



**HAL**  
open science

# An Eye Fixation-Related Potential Study in Two Reading Tasks: Reading to Memorize and Reading to Make a Decision

Aline Frey, Benoit Lemaire, Laurent Vercueil, Anne Guérin-Dugué

► **To cite this version:**

Aline Frey, Benoit Lemaire, Laurent Vercueil, Anne Guérin-Dugué. An Eye Fixation-Related Potential Study in Two Reading Tasks: Reading to Memorize and Reading to Make a Decision. *Brain Topography: a Journal of Cerebral Function and Dynamics*, 2018, 31 (4), pp.640 - 660. 10.1007/s10548-018-0629-8 . hal-01741895

**HAL Id: hal-01741895**

**<https://hal.science/hal-01741895>**

Submitted on 18 Jun 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# An Eye Fixation-Related Potential Study in Two Reading Tasks: Reading to Memorize and Reading to Make a Decision

Aline Frey<sup>1</sup> · Benoît Lemaire<sup>2</sup> · Laurent Vercueil<sup>3,4</sup> · Anne Guérin-Dugué<sup>5</sup>

Received: 13 September 2017 / Accepted: 5 January 2018  
© Springer Science+Business Media, LLC, part of Springer Nature 2018

## Abstract

We investigated how two different reading tasks, namely reading to memorize [Read & Memorize (RM)] and reading to decide whether a text was relevant to a given topic [Read & Decide (RD)], modulated both eye movements (EM) and brain activity. To this end, we set up an ecological paradigm using the eye fixation-related potentials (EFRP) technique, in which participants freely moved their eyes to process short paragraphs, while their electroencephalography (EEG) activity was recorded in synchronization with their EM. A general linear model was used to estimate at best EFRP, taking account of the overlap between adjacent potentials, and more precisely with the potential elicited at text onset, as well as saccadic potentials. Our results showed that EM patterns were top-down modulated by different task demands. More interestingly, in both tasks, we observed slow-wave potentials that gradually increased across the first eye fixations. These slow waves were larger in the RD task than in the RM task, specifically over the left hemisphere. These results suggest that the decision-making process during reading in the RD task engendered a greater memory load in working memory than that generated in a classic reading task. The significance of these findings is discussed in the light of recent theories and models of working memory processing.

**Keywords** Reading · Eye movements (EM) · Working memory (WM) · Eye fixation-related potentials (EFRP) · EEG · Slow wave potentials · General linear model (GLM)

## Introduction

Reading is a complex cognitive task involving an interaction between a reader and a text, for the purpose of constructing meaning from written representations. Over the past few decades, two online methodologies based on eye movements (EM) and event-related potentials (ERP) have led to a major breakthrough in understanding the processes and strategies

required in reading. There has recently been strong interest in coregistering EM and ERP, allowing for the study of eye fixation-related potentials (EFRP) to capitalize on the strengths of both techniques and overcome their separate weaknesses (Henderson et al. 2013).

In reading, this approach has several major benefits (Dimigen et al. 2011; Ditman et al. 2007; Frey et al. 2013; Henderson et al. 2013; Hutzler et al. 2007; Kliegl et al. 2012; Scharinger et al. 2015; Simola et al. 2009). Written linguistic symbols are spatially arranged, and because high-acuity vision is confined to a relatively small spatial area, readers need to coordinate their EM across the text in order to perform the task efficiently (comprehension, word searching, etc.; Weger and Inhoff 2007). The *eye-mind link* assumption suggests that the location of an observer's gaze reflects, at least in part, what is being processed in his or her mind at that time (Reichle and Reingold 2013). EM therefore constitute natural markers for segmenting the ongoing neural activity, and EFRP analysis allows for that activity to be investigated during self-paced perceptual and cognitive behaviors (Simola et al. 2015; Nikolaev et al. 2014). By using the temporal and spatial positions of specific eye

✉ Aline Frey  
aline.frey@u-pec.fr

<sup>1</sup> CHArt Laboratory, ESPE of the Créteil Academy, University of East-Paris Créteil, Site de Livry-Gargan, 45, avenue Jean Zay, 93190 Livry-Gargan, France

<sup>2</sup> Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, LPNC, 38000 Grenoble, France

<sup>3</sup> Univ. Grenoble Alpes, Inserm, CHU Grenoble Alpes, GIN, 38000 Grenoble, France

<sup>4</sup> EFSN, PPNR, CHU Grenoble Alpes, 38043 Grenoble, France

<sup>5</sup> Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, 38000 Grenoble, France

events (e.g., fixations or saccades) as markers for brain activity measures, the coregistration technique provides a means of gaining deeper insight into reading processes using more naturalistic experimental designs.

However, this technique still faces methodological issues (for a review, see Dimigen et al. 2011) which the present study was carried out to address. Specifically, given that fixations are generally short (about 250 ms), this results in many overlapping electroencephalography (EEG) responses. The mean latency of most of ERP components (e.g., P300, N400) is longer than the duration of a fixation plus a saccade. This means that the evoked response to a given fixation is partially superimposed on the previous and/or next one (Nikolaev et al. 2016; Ossandón et al. 2010). For example, the late components (e.g., N400) of a previous fixation may be superimposed on the early components (e.g., P100) of the current fixation. In addition, the potential elicited at stimulus onset overlaps with those elicited by the first fixations. Finally, oculomotor parameters can distort or mask the effects (e.g., saccade length or direction) of different experimental conditions. Nikolaev et al. (2016) emphasized this in the following example: "... suppose we are interested in the effect of stimulus size on visual information processing. Inspecting the larger stimulus, however, will require larger saccades. The EEG may differ solely because of larger saccades rather than because of any differences in information processing". A critical factor in EFRP analysis is thus the control of systematic differences in oculomotor variables (Simola et al. 2013). Linear models that break down the effects of overlap are a promising solution (Bardy et al. 2014; Burns et al. 2013; Congedo et al. 2016; Kristensen et al. 2017a; Smith and Kutas 2015a, b; Woldorff 1993). Estimation by regression with a general linear model (GLM) seems the most flexible method, as it has recently been implemented with success in EFRP/eye saccade-related potentials estimation (Dandekar et al. 2012; Kristensen et al. 2017b). Thus, in addition to a classic averaging method (Luck 2005), we also used a GLM method for EFRP estimation as an unbiased estimator with respect to the overlapping issue. Finally, and whichever estimation method is used, EFRP analysis also raises the problem of how to determine an adequate interval for baseline correction (Nikolaev et al. 2016). We therefore considered several possibilities and discussed their respective implications.

From a more theoretical point of view, relatively little is known about reading in more complex settings, as most research has been conducted using simpler protocols. And yet, finding and selecting relevant online information has become an important part of our everyday lives (Gordon and Pathak 1999; Pan et al. 2007). The Internet, in particular, can offer individuals new text formats, new purposes of reading, and new ways of interacting with information and texts that can have a major impact on their ability to approach and

understand a text and could lead to the emergence of specific cognitive processes (Leu et al. 2015). It is thus essential to investigate reading as part of online search tasks. To this end, we compared the effects of two different task demands on reading, using comparable reading material. The Read & Decide task (RD) simulates having to search the Internet for information about a specific topic. It involves two intertwined processes: reading and decision-making. Offering a reliable comparison, the Read & Memorize (RM) task entails reading with the intent of retaining the information that is encountered, and involves reading and memorization processes. It is akin to reading a recipe before cooking, which is a more careful reading than just getting the gist.

Ever since Yarbus's (1967) groundbreaking study of the visual exploration of paintings, it has been widely acknowledged that the nature of the task being performed influences EM. In reading, most studies have shown that EM responses are mainly and strongly modulated by the text's linguistic properties. A few studies have also shown that goal differences modulate EM patterns, and hence the selection of words in text reading (Kaakinen and Hyönä 2010; Radach et al. 2008). In Kaakinen and Hyönä's (2010) study, results showed that task instructions (i.e., proofreading or reading for comprehension) influence the initial landing position in words, as well as the saccade length, duration of the first fixation, and refixation probability. Moreover, Carver (1990) identified five reading strategies, based on variations in reading rates [words per minute (wpm)]. *Scanning* is performed at 600 wpm and is used, for example, when the reader is looking for a particular word contained in the text. *Skimming* (450 wpm) is typically observed in situations where the reader wants to gain an overview of the text without reading it in its entirety. *Reading* (300 wpm) corresponds to *normal* reading with the aim of understanding the text content. *Learning* (200 wpm) is a slow process used for acquiring new knowledge. Finally, *memorizing* (138 wpm) is the slowest process of all, involving continuous checks of whether previously encountered ideas are still remembered. Adopting the opposite approach, Simola et al. (2008) inferred hidden cognitive states, strongly related to different information search tasks, from observable EM behavior. Their model showed that EM provide the information needed to determine task type.

Complementary information provided by EEG measures of brain functioning can be used to compare different cognitive states and degrees of task engagement. In our study, we specifically focused on slow waves (i.e., deflections that typically last for > 200 ms and up to several seconds; Rösler 1993) as an indicator of the functioning of working memory (WM; Khader et al. 2005). Reading and WM interact closely with each other. As sentence or text processing extends over long time periods, readers have to maintain the thread of the linguistic input, as well as retrieve phonological, syntactic,

and semantic word properties from their long-term memory. Furthermore, this access to lexical information continuously interacts with and overlaps in time with the incremental integration of this information with the preceding context, to build up a message-level representation. The latter process refers to *unification*, which is the assembly of items stored in memory to form larger structures (Baggio and Hagoort 2011; Bastiaansen and Hagoort 2015) and is broadly sustained by WM (Hagoort 2013, 2016). The way that information in a text is combined with and integrated into the reader's knowledge has also been specifically expanded on in the well-known construction-integration model of text comprehension (Kintsch 1998; Weaver et al. 2012), in which memory plays an important part. Although the processes of integrating and constructing a coherent model of a text have been widely studied, and even implemented into computational versions (Kintsch 1998; Lemaire et al. 2006), much work remains to be done to understand how WM is used during reading. While there is general agreement that WM plays a fundamental role in text comprehension, there is also considerable disagreement over what WM is and what role it plays (Gordon et al. 2002).

Slow-wave topography and latencies have been shown to depend on the nature of the material (object, spatial or verbal stimuli) that has to be maintained in WM, and slow-wave amplitude varies according to the task's memory load and higher order attentional demands (Montfort and Pouthas 2003). Specifically, slow-wave amplitude recorded over frontal regions increases with WM demands, confirming the key role played by these frontal regions in the cognitive processes engaged during retention in WM. Most research on WM has used highly controlled laboratory paradigms (e.g., storage task or complex span task, i.e., a serial recall task where list presentations are interleaved with distracting subtasks). These are needed to characterize the basic function of WM, but are not very informative about how WM is used during a complex reading task (e.g., the positive-negative waveform, observed in an n-back task; Missonnier et al. 2003). Conversely, our approach to addressing this issue was to study WM involvement in ecological tasks where participants had to read several lines of text, instead of relying on non-ecological paradigms such as presenting one word at a time.

To our knowledge, only two studies have recently been undertaken to measure WM and/or cognitive load in unconstrained paradigms, using EFRP: Körner et al. (2014), using a visual search task, and Scharinger et al. (2015), using a reading task. In Körner et al.'s (2014) study, slow waves were identified during a two-target task, indicating WM involvement in the storage of spatial positions to prevent the revisiting of loci during the search. More specifically, participants saw a display featuring either one or two identical targets among several distractors, and were instructed

to indicate whether there were one or two targets in that display. Results showed a gradual build-up of a negativity shift from the first to the third distractor after the first target had been identified, indicating a change in brain activity most likely due to memory maintenance of the first target. For their part, Scharinger et al. (2015) constructed a protocol in which they induced additional load through hyperlink selection, comparing this with a baseline condition in which participants just had to read the text, with no hypertext processing. When they compared parts of the text that required pure reading (baseline condition) with parts of the text where participants had to interrupt their reading and perform hyperlink selection processes (test condition), they observed a significant decrease in alpha frequency band power, which they interpreted as an increase in memory load.

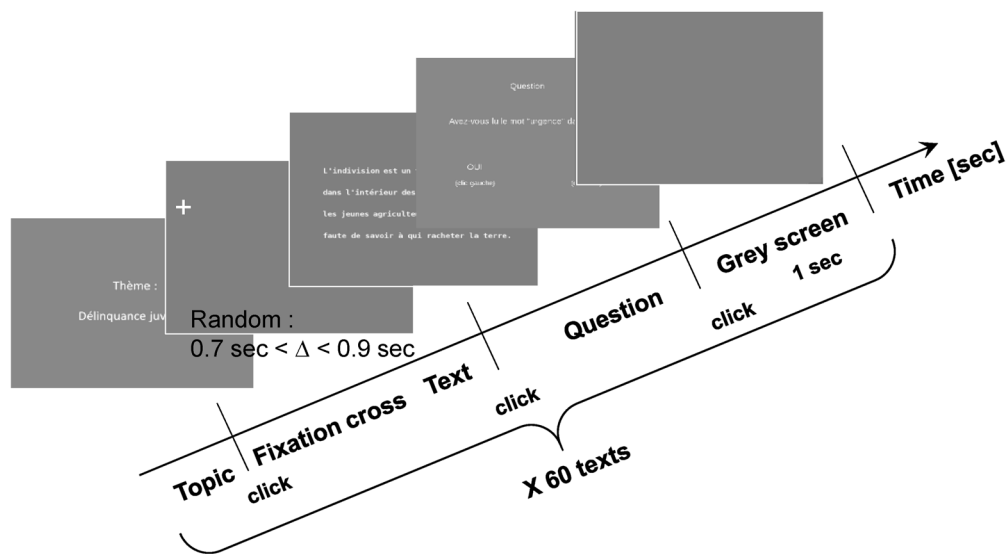
To sum up, despite a wide variety of research on brain responses to memory activation, we still do not understand exactly how WM is involved in everyday reading behavior. The goal of our study was thus to deepen our understanding of how EM sequences associated with neural markers are modulated together by administering the two reading tasks (RM and RD) mentioned above. Given the current state of knowledge, it was hard to predict which task (careful reading for memorizing or reading to quickly evaluate the semantic matching of the text content) would be more demanding with respect to WM.

## Materials and Methods

### Participants

Participants were 24 volunteers, all right-handed and French native speakers, with normal or corrected-to-normal visual acuity and no known neurological disorder. None of them had prior experience of the experimental task. They all gave their written informed consent prior to the experiment and were rewarded €20 in vouchers for their participation. Owing to technical problems with the ocular data (difficulty calibrating their eyes with satisfactory accuracy) and/or noisy EEG data (face/neck muscle activity, poor electrode contact), the data of four participants could either not be recorded or had to be excluded from the final analyses. The 20 remaining participants were ten women and ten men, aged 20–39 years ( $M = 27.2$  years,  $SD = 5$  years 2 months,  $SE = 1$  year 2 months).

The whole experiment was reviewed and approved by the ethics committee of Grenoble University Hospital (no. 2011-A00845-36). The EEG/eye-tracker coregistration was performed at the IRMaGe neurophysiology facility in Grenoble (France).



**Fig. 1** Trial timeline showing the different screens. Each condition (RM and RD) comprised 60 trials, each following this timeline

## Textual Material

We created 120 texts in French for the experimental session, and four for the familiarization session. These texts contained on average 1.8 sentences ( $SE = .13$ ), 5.18 lines ( $SE = .16$ ) and 30.1 words ( $SE = .65$ ). The mean length of each word was 5.34 letters ( $SE = .72$ ). Texts were printed in DejaVu font in white against a gray background.

For each text, we selected a topic (two or three words) that was either related or unrelated to the text. To assess the semantic relationship between a topic and a text, we used latent semantic analysis (Landauer et al. 2007). This method is based on the processing of a large sample of texts, in which a word or set of words (sentence, text, etc.) is represented by a vector in a 300-dimension semantic space. A cosine function is used to compute the similarity between the vectors. The higher the cosine value, the more similar the words or sets of words. In our case, we used a 24 million word corpus of French newspaper articles. For all related topics, semantic similarity with the text was above 0.2, while for all unrelated topics, semantic similarity was below 0.06.

## Experimental Procedure

Every participant completed both tasks in the same order (RM then RD). For each participant, the 120 texts were randomly split into two blocks (60 texts per task) in such a way that each text was presented the same number of times in each condition across the whole experiment. For each trial, participants were provided with a topic, a text, and a question they had to answer. In the RD task, participants were asked to imagine they were doing a press review and had to decide whether to retain or reject a text depending on

the topic they had been given (Q: “Do you want to retain or reject the text?”; with two possible answers that were counterbalanced), as quickly as possible during their reading (as emphasized by Scharinger et al. (2015), it is also important to construct realistic reading situations, hence the time limit for text reading in our study). Participants were told that the answer “Yes” meant that the text was semantically related to the topic, and the answer “No” meant that it was not. In the RM task, participants were asked to undertake a careful reading of the text in order to memorize the words. They were then asked to answer one of three questions. Two questions about the presence or otherwise of a word in either the topic or the text (Q1: “Did you see the word ‘xxxx’ in the topic?”; Q2: “Did you see the word ‘yyyy’ in the text?”). We also wanted to encourage the creation of a semantic link between the topic and the text in the RM task, to make sure that both tasks were similar in this regard. To this end, participants were additionally asked to answer a question about the semantic link (Q3: “Did the topic help you to understand the text?”). These three questions were counterbalanced across participants, ensuring that there were as many positive as negative answers.

The timeline of each trial is illustrated in Fig. 1. Each trial started with the presentation of the topic, followed by the display of a fixation cross, then the text, the question, and finally a gray screen before the next trial. The fixation cross was displayed to the left of the first letter of the first word of the text, for a random duration of 2100–2500 ms, to avoid an anticipatory saccade before the text appeared. The text was displayed once the participant’s gaze had stabilized on the fixation cross (in a rectangle of  $3^\circ \times 2^\circ$  pixels around the fixation cross). Participants had to read the topic, read the text, answer the question, then click on the mouse button



(with their dominant hand) when they were ready to go to the next step.

In both conditions, a pause was scheduled after 60 trials, and participants were given the opportunity to rest if necessary. Before the experimental phase, four trials (two for each condition) were run to familiarize participants with the experimental design. The experiment was built using SoftEye software (Ionescu et al. 2009) to control the timing of the displays, the eye-tracker, and the sending of synchronization triggers to both devices (i.e., EEG and eye-tracker).

## Data Recording

The whole experiment lasted about 90 min, during which participants were comfortably seated in an adjustable chair, in a sound-attenuated and dimly lit room.

### Eye Movement Data Recording

EM were recorded with a video-based infrared eye-tracking system (EyeLink® 1000; SR Research Ltd., Ontario, Canada), in pupil-corneal reflection tracking mode, with a sampling rate of 1000 Hz. Participants' head position was stabilized with a chin rest. Only the dominant eye was recorded.

Participants were seated 57 cm from a 20-in. monitor (38° × 33° visual field) with a screen resolution of 1024 × 768 pixels. The text was displayed in the center of the screen (26° × 9° visual field), with 50.7 letters per line on average. Thus, each letter covered a horizontal visual angle of 0.46°, corresponding to about 4.3 letters in the foveal region.

At the beginning of the experiment, we carried out a 9-point calibration routine. Drift correction was performed every three trials, and calibration every 15 trials, ensuring accurate recording of the eye's fixations on the screen. Calibration was automatically initiated if an offset above 0.5° was detected, and the eye-tracker could also be recalibrated if the experimenter deemed it necessary.

### EEG Data Recording

EEG activity was recorded using a 10-20 system (Jasper 1958) ActiCap containing 64 channels (BrainCap™; Brain Products GmbH, Gilching, Germany). The impedance of the electrodes was kept below 5 kΩ by using contact gel (SuperVisc Gel, Brain Products, Inc.) to increase the signal-to-noise ratio. Electrodes were referenced to the head (FCz; ground: AFz) during the recording, and re-referenced offline to the mean mastoid. EEG data were amplified using BrainAmp™ (Brain Products GmbH), sampled at 1000 Hz.

Four EOG electrodes were affixed to the outer canthi and infraorbital ridges of both eyes. The seating position and

head position in the eye-tracker were carefully adapted to avoid myogenic artifacts from neck and temple muscles.

## Data Preprocessing and Analysis

### Eye Movement and EEG Data Integration

Both the EEG and eye-tracker data were sampled at the same rate (1000 Hz). EM and EEG signals were synchronized offline, on the basis of triggers sent simultaneously during the experiment to both the EEG system and the eye-tracker.

### Eye Movement Analysis

Saccades were automatically detected from the raw eye coordinate data using the standard saccade detection algorithm with the following thresholds: minimum velocity 30°/s, minimum acceleration 8000/s/s, and minimum motion 0.15°. A fixation was defined here as the maintenance of the visual gaze on a single location that was neither a blink nor a saccade. In line with previous experiments (Dimigen et al. 2011), shorter and longer fixations (i.e., fixations lasting < 80 ms or > 1000 ms) were excluded. For each trial, we considered the first fixation (Rank 1) as the one starting after text onset. The last one (Rank N) was the one ending before text offset.

### EEG Data Preprocessing

EEG data preprocessing and EFRP analyses were carried out with BrainVision Analyzer 2 software (Version 2.0.4; Brain Products GmbH, Munich, Germany). Continuous EEG was first segmented across the whole duration of each text. An analog bandpass filter (0.01–100 Hz) was applied, as well as a 50-Hz notch filter (symmetrical 5-Hz bandwidth around the notch frequency, i.e., 50 ± 2.5 Hz; 24 dB/oct) to eliminate interference from mains electricity. Segments containing artifacts (excessive max–min) were rejected using a semi-automatic artifact rejection procedure (55.7 (4.0) remaining trials on average in the RM task and 55.15 (3.1) in the RD). Segments were then broken down by infomax independent component analysis (ICA; Jung et al. 2000; Makeig et al. 1997), and ICA components reflecting eye blinks and saccades were removed. Finally, trials with fewer than seven fixations were excluded for methodological and theoretical reasons: first, to make sure that the EFRP results were not contaminated with ERP related to response programming and execution (van Vugt et al. 2014); and second, to eliminate excessively brief trials to allow us to observe a reading process in both tasks. The number of trials was then equalized for each participant across the tasks, by randomly removing trials from the RM task. On average, there were 42.7 (7.7) remaining trials in each of the two tasks. After

this selection, the minimum number of fixations in the RM task was 14.

Six regions of interest (ROI), each grouping four electrodes, were defined as follows: frontal right (F4, F6, F8, FC6); frontal left (F3, F5, F7, FC5); central right (C4, C6, T8, CP6); central left (C3, C5, T7, CP5); parietal right (P2, P4, P6, PO4); and parietal left (P1, P3, P5, PO3).

## Results and Models

### Behavioral Results

#### Response Rate

The rate of correct answers was 97.8% ( $SE = .49$ ) for the RD task and 78.4% ( $SE = 1.88$ ), for the RM task,  $t(19) = -10.39$ ,  $p < .001$ ,  $\eta^2 = .86$ . More specifically, there were 84.1% correct answers ( $SE = 3.33$ ) for the question about the presence or absence of a word in the topic, and 72.5% ( $SE = 3.55$ ) for the question about the presence or absence of a word in the text.

#### Reading and Response Times

Regarding the amount of time spent reading the text, results showed a significant difference,  $t(19) = 15.53$ ,  $p < .001$ ,  $\eta^2 = .93$ , between the two conditions: participants needed more time to read the texts in the RM condition (10.05 s;  $SE = .58$ ) than in the RD condition (3.17 s;  $SE = .25$ ). Results also showed a significant difference between the two conditions regarding the time taken to answer the question,  $t(19) = 17.66$ ,  $p < .001$ ,  $\eta^2 = .95$ , with a mean time of 2.95 s ( $SE = .16$ ) in the RM condition, and 0.71 s ( $SE = .04$ ) in the RD condition.

### Eye Movement Results

#### Reading Behavior

The following EM variables were used to assess the extent to which readers' behavior was affected by the nature of the task: number and mean duration of fixations, duration of inter-fixation interval (IFI<sup>1</sup>), reading speed, percentage of refixations (consecutive fixations on the same word), and percentage of long progressive fixations (consecutive fixations skipping words; Underwood and Radach 1998).

In line with the text reading times (see above, “Behavioral Results” section), the EM analysis across participants revealed that there were more fixations in the RM condition (37.66;  $SE = 1.92$ ) than in the RD condition (12.31;  $SE = .83$ ), and this difference was significant,  $t(19) = 17.77$ ,  $p < .001$ ,  $\eta^2 = .95$ . Moreover, mean fixation duration was significantly higher in the RM condition (214.44 ms;  $SE = 5.30$ ) than in the RD condition (185.53 ms;  $SE = 6.08$ ),  $t(19) = 6.96$ ,  $p < .001$ ,  $\eta^2 = .73$ . Mean IFI duration was significantly higher in the RM condition (254.58 ms;  $SE = 5.28$ ), than in the RD condition (227.03 ms;  $SE = 5.86$ ),  $t(19) = 6.45$ ,  $p < .001$ ,  $\eta^2 = .70$ . We then computed the mean reading speed for each participant as the mean number of words read per minute (wpm), for all texts. Specifically, we calculated the reading speed by considering the position of the word corresponding to the farthest fixated word, relative to the text reading time. Mean reading speed in the RM condition was 199.00 wpm ( $SE = 12.80$ ), which was significantly lower than in the RD condition (488.15 wpm,  $SE = 25.27$ ),  $t(19) = -11.22$ ,  $p < .001$ ,  $\eta^2 = .87$ .

The percentage of refixations (when there were two consecutive fixations on the same word, the second one was counted as a refixation) was greater for the RM task (7.74%;  $SE = .87$ ) than for the RD task (2.64%;  $SE = .51$ ) and this difference was significant,  $t(19) = 8.23$ ,  $p < .001$ ,  $\eta^2 = .79$ .

Long progressive fixations (i.e., where at least one whole word was skipped) were also analyzed. The percentage of long progressive fixations was significantly lower for the RM task (5.25%;  $SE = .54$ ) than for the RD task (14.32%;  $SE = 1.03$ ),  $t(19) = -10.30$ ,  $p < .001$ ,  $\eta^2 = .85$ .

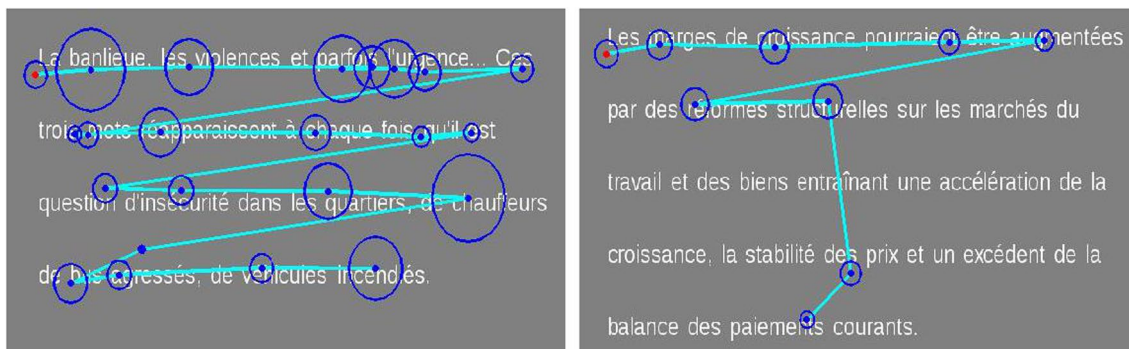
Figure 2 presents an example of a scanpath for each task, showing the differences between the two tasks regarding the number of fixations, fixation duration, number of refixations and number of long progressive fixations.

#### Fixations of Interest for EFRP Analyses

Table 1 summarizes the mean duration ( $SE$ ) and median duration ( $SE$ ) of the fixations of interest selected for the EFRP analyses (i.e., first four fixations) for all the trials and participants, for each task.

An analysis of variance (ANOVA) including the factors condition and fixation rank showed a main effect of condition for mean duration,  $F(1, 19) = 41.27$ ,  $p < .001$ ,  $\eta^2 = .68$ , and for median duration,  $F(1, 19) = 24.29$ ,  $p < .001$ ,  $\eta^2 = .56$ , with a fixation duration that was significantly lower for the RD task (mean = 188.04, median = 180.84) than for the RM task (mean = 217.03, median = 203.00). The main effect of rank was also significant for mean duration,  $F(3, 57) = 22.22$ ,  $p < .001$ ,  $\eta^2 = .52$ , and for median duration  $F(3, 57) = 20.20$ ,  $p < .001$ ,  $\eta^2 = .52$ , as the first fixation (mean = 185.41, median = 176.12) was shorter than the other three (all  $p < .001$ ). The difference between the

<sup>1</sup> The interval between the offset of a fixation and the onset of the following one (thus including the duration of the fixation plus the saccade).



**Fig. 2** Example of scanpaths for the RM task (left) and RD task (right). The center of the circle is the fixation position and its radius is proportional to the fixation duration

**Table 1** Statistical summary of the fixations of interest for EFRP estimations based on individual means

Fixations of interest Condition	Mean fixation duration (SE) in ms		Median fixation duration (SE) in ms	
	RM	RD	RM	RD
Fix. 1	201.91 (6.98)	168.92 (6.03)	191.47 (6.44)	160.78 (4.70)
Fix. 2	218.93 (6.32)	193.00 (7.75)	203.70 (5.90)	185.90 (7.91)
Fix. 3	225.30 (6.11)	197.69 (6.07)	209.28 (5.56)	189.28 (6.16)
Fix. 4	221.99 (5.48)	192.55 (6.98)	207.55 (5.61)	187.40 (6.81)

tasks on fixation duration for all the fixations on the text (see above, “**Reading Behavior**” section) was also observed for the four fixations of interest in EEG analysis.

Concerning the first four fixations of interest, the percentages of long progressive fixations and refixations differed according to the task. As with the whole text, the percentage of refixations was greater for the RM task (3.79%,  $SE = .51$ ) than for the RD task (0.92%,  $SE = .20$ ), and this difference was significant,  $t(19) = 6.72$ ,  $p < .001$ ,  $\eta^2 = .71$ . The percentage of long progressive fixations was significantly lower for the RM task (1.56%,  $SE = .45$ ) than for the RD task (4.74%,  $SE = .91$ ),  $t(19) = -4.36$ ,  $p < .001$ ,  $\eta^2 = .51$ . These differences were re-investigated from low-level oculomotor features, by evaluating the amplitude and direction of the incoming saccades. Considering the four fixations of interest in each task, we divided the incoming saccades into four categories according to two criteria: direction (progressive vs. regressive) and amplitude (short vs. long; split according to the median value ( $3.46^\circ$ ) whatever the direction and the task). The sizes of the four categories for each task were submitted to an ANOVA including the condition (RM vs. RD), direction (progressive vs. regressive) and amplitude (short vs. long) factors. As expected, the main effect of condition was not significant (see “**EEG Data Preprocessing**” section; the number of trials was then equalized for each participant across the tasks), nor was the main effect of amplitude. The main effect of direction was significant,  $F(1, 19) = 271.9$ ,  $p < .001$ ,  $\eta^2 = .93$ , with more progressive saccades (82.25) than regressive ones (12.46). Results also

showed a Condition x Direction interaction,  $F(1, 19) = 10.51$ ,  $p < .005$ ,  $\eta^2 = .36$ , and a Condition x Amplitude interaction,  $F(1, 19) = 57.26$ ,  $p < .001$ ,  $\eta^2 = 0.75$ . Finally, the Condition x Direction x Amplitude interaction was significant,  $F(1, 19) = 22.84$ ,  $p < .001$ ,  $\eta^2 = .55$ . Post hoc comparisons (Tukey test) showed that there were more short progressive saccades in the RM condition (85.5) than in the RD one (62.5,  $p < .001$ ), and conversely, more long progressive saccades in the RD condition (108) than in the RM one (73,  $p < .001$ ). No significant differences between the two conditions were observed regarding regressive saccades (short vs. long).

### EEG Data Analysis by Averaging

In order to capture the growth of slow waves, we were first inspired by Körner et al.’s (2014) methodology, and applied the classic averaging method (Luck 2005), time-locking the EEG signal to the onset of each of the first four fixations (i.e., fixations of interest) for each text.

The choice of an appropriate baseline in EFRP analysis is always very challenging (see Luck 2005; Dimigen et al. 2011; Körner et al. 2014; Nikolaev et al. 2016). An individual baseline, related to each fixation of interest, is mainly used to investigate fast processes specifically related to the current fixation, based on the premise that these processes may change from one fixation to another. This baseline has the advantage of being temporally close to the events being analyzed and providing a common reference point for



comparing the selected fixations (Körner et al. 2014). On the other hand, a *common* baseline “is applied to all saccade related epochs within one trial. This type of baseline allows investigation of accumulative processes occurring across sequential EM, such as memory buildup, changes in attention and effort” (Nikolaev et al. 2016, p. 66). As we aimed to capture slow waves, that is, deflections that typically unfold over longer durations (i.e., > 200 ms and up to several seconds), it seemed obvious to us that we should use a *common* baseline. The overriding criterion of choice in the EFRP literature is to avoid any corruption of the baseline by preceding and following fixations. Thus, for the four fixations of interest, we used a common –100 to 0 ms interval before text onset (as in classic ERP studies; for a review, see Woodman 2010), during which we would not expect the brain response to differ between trials. As a control, and to confirm that expected slow waves were not synchronized at the onset of each fixation, we also conducted analyses with an *individual* baseline, taking a segment of –200 to –100 ms before each fixation onset, to avoid corruption by pre-saccadic activity. The detailed results of these individual baseline analyses are provided in “Appendix 1”.

### Estimation by averaging

Thus, the signal  $x_i^{(r)}(t)$  time-locked to the  $r$ th fixation onset during the  $i$ th epoch can be written as:

$$x_i^{(r)}(t) = fp^{(r)}(t) + n_i(t) \quad (1)$$

where  $fp^{(r)}(t)$  is the potential evoked at the  $r$ th fixation onset and  $n_i(t)$  is the background cortical activity, considered as noise. Assuming that all fixations at the  $r$ th rank elicited the same potential and that the ongoing activity was not synchronized with the fixation onset during the  $i$ th epoch, this potential was estimated by averaging on a given number of epochs, as follows:

$$\widehat{fp}_{Av}^{(r)}(t) = \frac{1}{E} \sum_{i=1}^E x_i^{(r)}(t). \quad (2)$$

This estimator is only unbiased if a single potential is elicited per epoch (Ruchkin 1965), and to take our analyses further, we adopted an alternative solution using the GLM described below.

To minimize EM effects on the EEG signal (especially at the onset of the subsequent saccade), we restricted the analysis window duration to a period during which the eyes were not moving (Körner et al. 2014). For each of the four fixations of interest, we segmented 150-ms epochs beginning at the onset of fixation. This duration was chosen to

be shorter than the shortest median<sup>2</sup> fixation duration value (i.e., 160 ms; see Table 1), to best ensure that the eyes were fixating.

In both analyses (common and individual baselines), the mean amplitude values within the 150-ms window were entered in a repeated-measures ANOVA with the condition (RM, RD), hemisphere (left, right), ROI (frontal, central, parietal) and rank (1, 2, 3, 4) factors.  $P$  values were reported after the Greenhouse–Geisser correction for nonsphericity (Greenhouse and Geisser 1959) and Tukey tests were used for post hoc comparisons.

### Results on EFRP by Averaging

Based on the potentials elicited at the first four fixation onsets, estimated by averaging, the ANOVA on the mean amplitudes with a common baseline revealed the following significant effects:

Condition:  $F(1, 19) = 10.26, p < .005, \eta^2 = .35$

Condition  $\times$  Hemisphere interaction:  $F(1, 19) = 5.4, p < .05, \eta^2 = .22$

Condition  $\times$  ROI interaction:  $F(2, 38) = 3.71, p = .05, \eta^2 = .16$

Condition  $\times$  Rank interaction:  $F(3, 57) = 14.77, p < .001, \eta^2 = .44$

Condition  $\times$  Hemisphere  $\times$  ROI interaction:  $F(2, 38) = 7.11, p = .01, \eta^2 = .27$

Condition  $\times$  Hemisphere  $\times$  Rank interaction:  $F(3, 57) = 7.27, p = .01, \eta^2 = .28$

Condition  $\times$  Hemisphere  $\times$  ROI  $\times$  Rank interaction:  $F(6, 114) = 6.64, p = .01, \eta^2 = .26$

To avoid redundancy, we only provide the results of the post hoc comparison for the Condition  $\times$  Hemisphere  $\times$  ROI  $\times$  Rank interaction. Moreover, we present the results for the frontal, central and parietal ROI separately (Table 2).

Over the left hemisphere (frontal ROI; i.e., F3, F5, F7, FC5), the post hoc comparison revealed that in the RM and RD conditions, negativity significantly increased across the four fixations, with the first fixation eliciting significantly less negativity than the second one, which in turn elicited less negativity than the third one, which was also less negative than the fourth one (Fig. 3).

Over the right hemisphere (frontal ROI; i.e., F4, F6, F8, FC6), in the RM condition, the first three fixations did not differ from each other, and Fixations 1 and 2 elicited less positivity than Fixation 4. In the RD condition, the second

<sup>2</sup> A median value was a more suitable indicator than a mean value, owing to the asymmetrical distribution.

**Table 2** Mean (standard error) amplitude values in  $\mu\text{V}$  for the latency ranges of interest for frontal, central and parietal ROI in each condition with a common baseline

		Rank 1	Rank 2	Rank 3	Rank 4	Significant differences
<b>Frontal ROI amplitude [<math>\mu\text{V}</math>]</b>						
Left Hem.	RM	1.84 (.92)	-1.29 (.91)	-3.45 (1.01)	-4.84 (1.16)	1 > 2 > 3 > 4**
Left Hem.	RD	-.10 (.77)	-7.21 (1.22)	-11.95 (1.59)	-14.57 (1.57)	1 > 2 > 3 > 4***
Significant differences		RM > RD ***	RM > RD ***	RM > RD ***	RM > RD ***	
Right Hem.	RM	4.27 (.99)	4.57 (1.09)	5.51 (1.04)	6.48 (1.15)	1 = 2 = 3 & (1 = 2) < 4***
Right Hem.	RD	3.70 (1.09)	2.11 (1.43)	3.49 (1.77)	3.86 (1.77)	2 < (1 = 3 = 4)*
Significant differences		ns	RM > RD ***	RM > RD ***	RM > RD ***	
<b>Central ROI amplitude [<math>\mu\text{V}</math>]</b>						
Left Hem.	RM	3.59 (.91)	1.68 (.89)	.48 (.80)	-.27 (.84)	1 > 2 > (3 = 4) & (2 > 4)
Left Hem.	RD	2.32 (.86)	-2.76 (1.03)	-5.70 (1.13)	-7.59 (.89)	1 > 2 > 3 > 4***
Significant differences		RM > RD *	RM > RD ***	RM > RD ***	RM > RD ***	
Right Hem.	RM	4.43 (.95)	3.86 (.91)	3.72 (.84)	3.99 (.91)	ns
Right Hem.	RD	3.19 (1.07)	.59 (1.29)	.12 (1.47)	-.57 (1.30)	1 > (2 = 3 = 4)***
Significant differences		ns	RM > RD ***	RM > RD ***	RM > RD ***	
<b>Parietal ROI amplitude [<math>\mu\text{V}</math>]</b>						
Left Hem.	RM	5.05 (.96)	2.90 (.99)	1.54 (.90)	1.27 (.92)	1 > 2 > 3 = 4*
Left Hem.	RD	4.40 (1.04)	0.00 (1.16)	-2.72 (1.23)	-4.39 (1.00)	1 > 2 > 3 > 4***
Significant differences		ns	RM > RD ***	RM > RD ***	RM > RD ***	
Right Hem.	RM	5.34 (1.01)	3.57 (.95)	2.40 (.85)	2.32 (.86)	1 > (2 = 3 = 4)
Right Hem.	RD	4.44 (1.18)	.68 (1.30)	-1.33 (1.38)	-2.62 (1.13)	1 > 2 > 3 > 4*
Significant differences		ns	RM > RD ***	RM > RD ***	RM > RD ***	

Significant differences in the rows correspond to post hoc analyses between the four ranks of fixations, and those in the columns correspond to post hoc analyses between the two tasks. \* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$

fixation elicited less positivity than the other three fixations, which did not differ from each other.

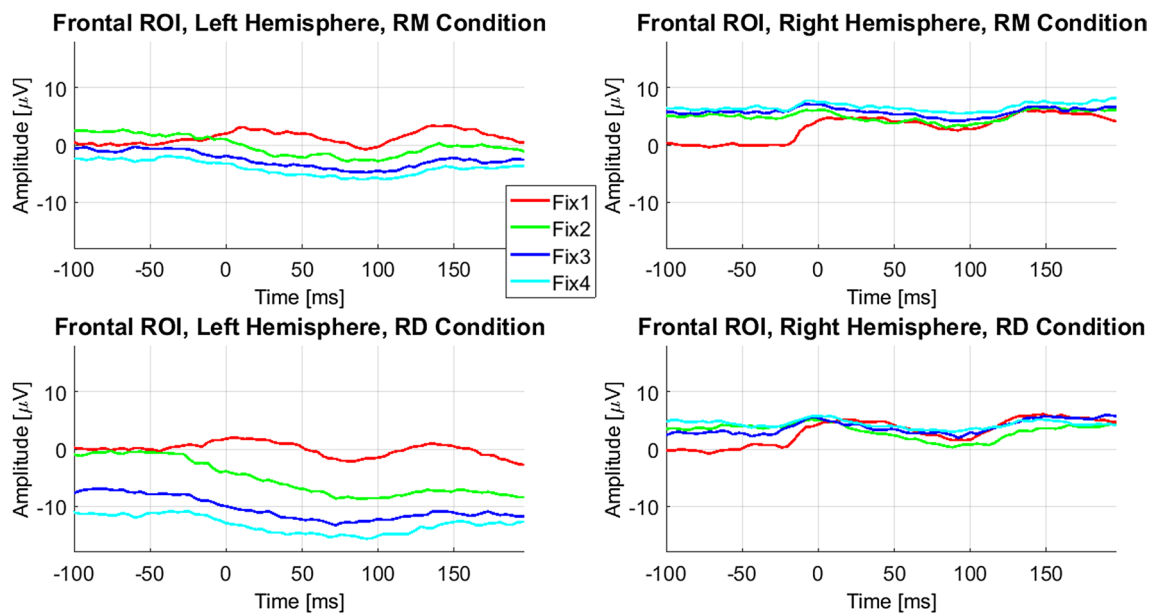
Moreover, over the left hemisphere, the four fixations elicited significantly greater negativity in the RD condition than their equivalent in the RM condition. Over the right hemisphere, Fixations 2, 3, and 4 elicited significantly greater negativity in the RD condition than their equivalent in the RM condition.

Figures 4 and 5 illustrate the EFRP results for the central ROI (i.e., C4, C6, T8 and CP6 for the right hemisphere, and C3, C5, T7 and CP5 for the left one) and parietal ROI (i.e., P2, P4, P6 and PO4 for the right hemisphere, and P1, P3, P5 and PO3 for the left one).

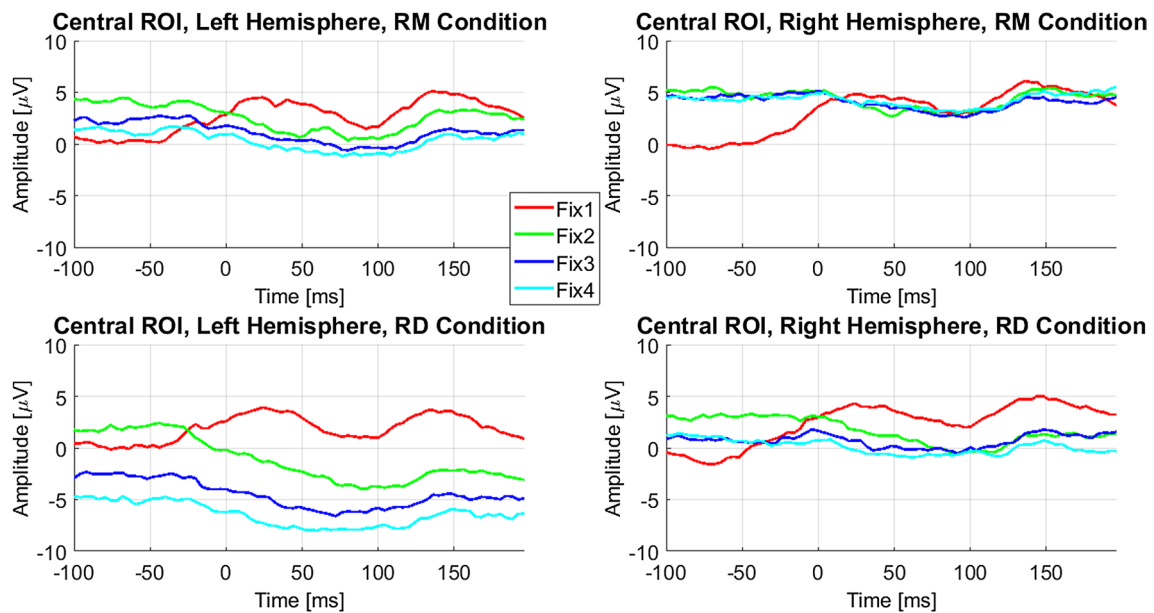
For the central left ROI in the RM condition, Fixation 1 elicited less negativity than the other three. Fixations 3 and 4 did not differ from each other, while Fixation 2 also elicited less negativity than Fixation 4. In the RD condition, negativity increased significantly across the four fixations (Fixation 1 elicited significantly less negativity than Fixation 2, which in turn elicited less negativity than Fixation

3, which was also less negative than Fixation 4). Over the right hemisphere, no significant difference was observed in the RM condition. In the RD condition, Fixations 2, 3 and 4 did not differ from each other and were significantly less negative than Fixation 1. Moreover, over the left hemisphere, the four fixations elicited significantly greater negativity in the RD condition than their equivalent in the RM condition. Over the right hemisphere, Fixations 2, 3, and 4 elicited significantly greater negativity in the RD condition than their equivalent in the RM condition (Fig. 4).

For the parietal left ROI in the RM condition, Fixation 1 elicited less negativity than Fixation 2, which in turn elicited less negativity than Fixation 3. Fixations 3 and 4 did not differ from each other. In the RD condition, negativity significantly increased across the four fixations (Fixation 1 elicited significantly less negativity than Fixation 2, which in turn elicited less negativity than Fixation 3, which was also less negative than Fixation 4). Over the right hemisphere, in the RM condition, Fixations 2, 3 and 4 did not differ from each other and were significantly more negative than Fixation 1.



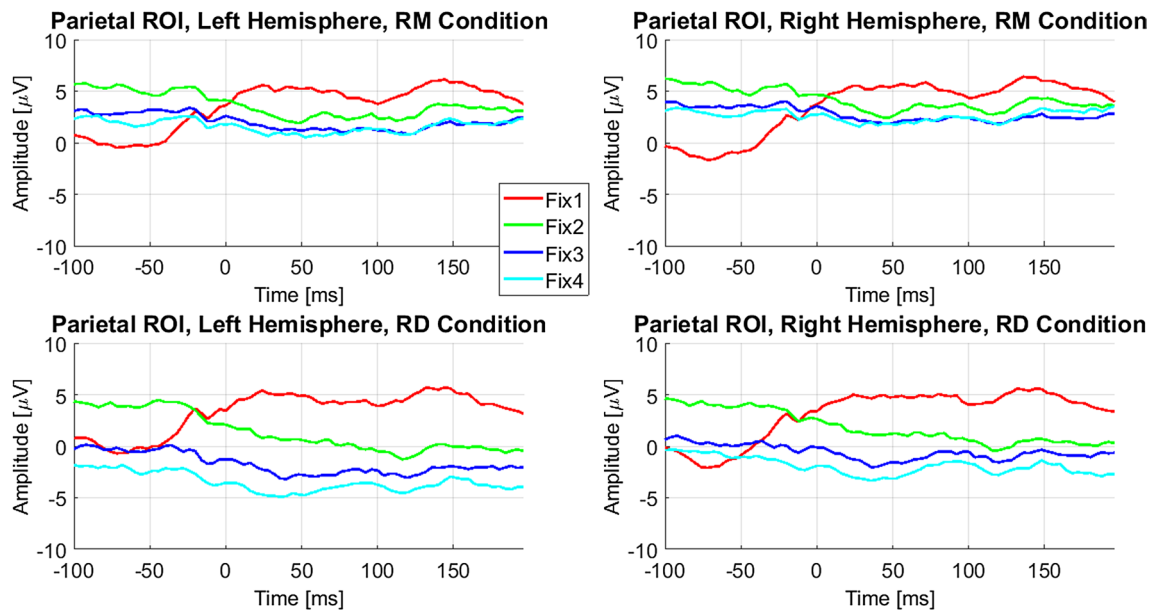
**Fig. 3** Grand-average EFRP, estimated by averaging with a common baseline, for the RM task (top) and RD task (bottom), for the first four fixations on the text, for the frontal right (right) and frontal left (left) ROI



**Fig. 4** Grand-average EFRP, estimated by averaging with a common baseline, for the RM task (top) and RD task (bottom), for the first four fixations on the text, for the central right (right) and left (left) ROI

In the RD condition, negativity increased significantly across the four fixations (Fixation 1 elicited significantly less negativity than Fixation 2, which in turn elicited less negativity than Fixation 3, which was also less negative than Fixation 4). Moreover, over both hemispheres, Fixations 2, 3, and 4 elicited significantly greater negativity in the RD condition than their equivalent in the RM condition.

A similar analysis, with the same estimates but with an individual baseline, was also conducted. Data and figures yielded by the ANOVA on the mean amplitudes are provided in “[Appendix 1](#)”. We therefore set out only the main findings here. Our results showed no major increment or decrement in synchronized activity elicited at one fixation compared with the following fixation, such that the level of



**Fig. 5** Grand-average EFRPs, estimated by averaging with a common baseline, for the RM task (top) and RD task (bottom), for the first four fixations on the text, for the parietal right (right) and left (left) ROI

synchronized activity at Fixation  $n$  was roughly the same as that at Fixation  $n + 1$ . As expected, slow-wave activity was not synchronized with each fixation.

### EEG Data Analysis by Regression

The results of the analysis by averaging provided an initial overview of the slow waves generated by each task. With a *common* baseline, results showed that amplitudes followed a gradient across successive fixations, but this outcome needed to be confirmed using a more suitable estimation method. The major concern with the EFRP estimation by averaging was the overlap with the potential elicited at text onset. This lasted about 700 ms, corresponding to the time needed for the stimulus-evoked activity to fade (Dimigen et al. 2011; Nikolaev et al. 2016). Moreover, as each task was different, the hypothesis of the same evoked potential could appear questionable. This overlap became smaller across fixations. Hence, the first potential was more impacted than the second one, and so on. To compare the slow waves generated by each task in the most equivalent way possible, it was thus very important to clean up the ERP elicited by the text. Also, as we were interested in neural activity over frontal sites, and as we observed significant differences in the amplitudes and directions of incoming saccades between the two tasks (see “Fixations of Interest for EFRP Analyses” section), we estimated saccade-related potentials to remove them from the estimation of the potential of interest. In the following two sections, we first present the simple regression model

(GLM) based on the previous results, then the results for the slow waves, characterized directly by their slope in  $\mu\text{V}/\text{ms}$ .

### Estimation by Regression

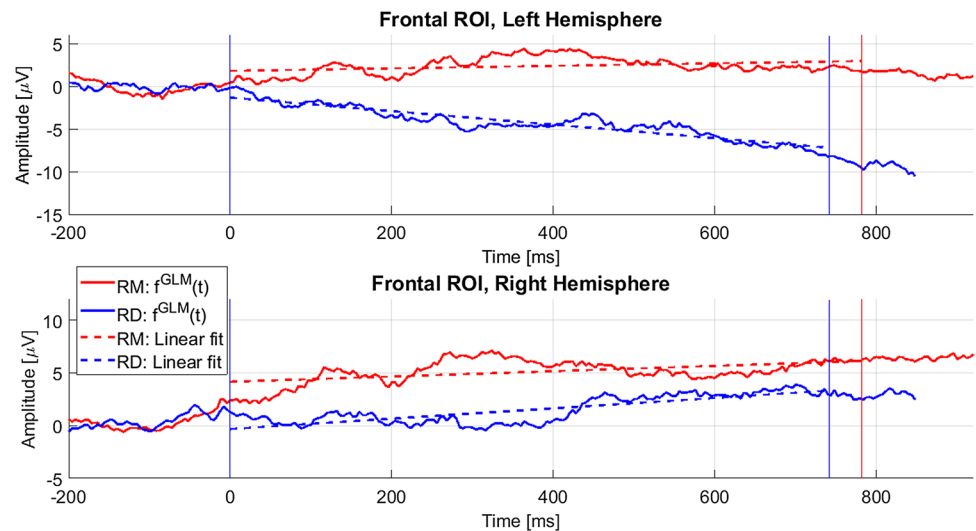
Analyses with an individual baseline showed that amplitudes did not follow a gradient across fixations, suggesting that the slow waves were not directly time-locked to each fixation onset. More specifically, we observed slow-wave neural activity time-locked to first fixation onset. This neural activity was estimated by a GLM, taking account of overlapping responses (evoked potential at text onset and saccadic potentials) to provide an unbiased estimator. Assuming that the slow waves were not evoked by a specific fixation, but rather continually fueled during task resolution, each signal  $x_i(t)$  time-locked to text onset during the  $i$ th epoch<sup>3</sup> and including the first four fixations was modeled by the following equation:

$$x_i(t) = s(t) + fp(t - \tau_i^{(1)}) + \sum_{c=1}^4 \sum_l sp_c(t - \tau_{c,i}^{(l)}) + n_i(t) \quad (3)$$

where  $s(t)$  is the evoked potential at text onset,  $fp(t)$  is the potential evoked at first fixation onset *as well as* the potentials elicited at subsequent fixation onsets up to the end of the epoch,  $sp_c(t)$  is the potential evoked at saccade

<sup>3</sup> Each epoch started 100 ms before text onset and ended 1050 (RD) or 1150 (RM) ms after text offset (see “Appendix 2”).

**Fig. 6** Grand averages of the first GLM-estimated EFRP  $\hat{f}p_{GLM}(t)$  for the frontal left (top) and right (bottom) ROI. The linear fitting is plotted in dotted line



onset for a given category ( $c$ ) of incoming saccade, and  $n_i(t)$  is the noise of the ongoing activity. Regarding the statistical distribution of the features (direction and amplitude) of the incoming saccades associated with the fixations of interest, four categories of incoming saccades were defined (see Subsection 3.2.2 “Fixations of interest for EFRP analyses”). In this equation, for a given epoch ( $i$ ),  $\tau_i^{(1)}$  is the timestamp of the first fixation onset, and  $\tau_{c,i}^{(l)}$  is the timestamp of an incoming saccade (rank  $l$ ) belonging to the  $c$ th category. This model was designed to observe the slow wave resulting from a continuous process from first fixation onset.

The interpretation of the potential  $fp(t)$  depends on  $t$ . From first fixation onset to first fixation offset (i.e.,  $t$  greater than zero and less than the duration of the first fixation<sup>4</sup>), the waveform  $fp(t)$  indicates that activity was only synchronized with the first fixation onset, that is, the first EFRP. However, for larger  $t$ , the waveform  $fp(t)$  indicates that activity was less and less synchronized with the first fixation onset. In other words, as  $t$  increased, the potential  $fp(t)$  gradually integrated the activity evoked by subsequent fixations. Taken together, for larger  $t$ ,  $fp(t)$  reveals averaged activity elicited by successive fixations, and only coherent deviations such as slow waves can emerge above the noise level across the epochs.

By concatenating all the trials, we could estimate  $s(t)$  and  $fp(t)$  using ordinary least square regression. Mathematical details for implementing the GLM are given in “Appendix 2”, along with all the configuration parameters for these estimations. As in the analysis by averaging, the model was applied separately to obtain the estimated  $\hat{f}p_{GLM}(t)$  for each participant. The slow waves for each electrode cluster (ROI)

were then characterized by their slope, which was computed from a linear fitting on the estimated  $\hat{f}p_{GLM}(t)$ . The slope values were entered in a repeated-measures ANOVA, with the condition (RM, RD), hemisphere (left, right) and ROI (frontal, central, parietal) factors.  $P$  values were reported after Greenhouse–Geisser correction for nonsphericity (Greenhouse and Geisser 1959) and Tukey tests were used for post hoc comparisons.

### Results for EFRP by Regression

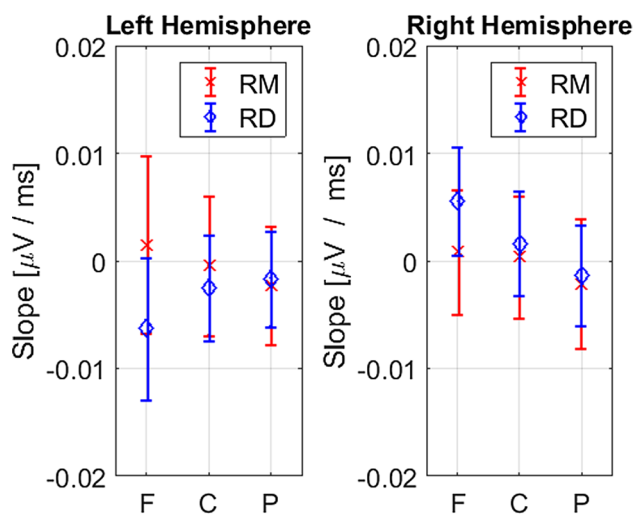
Figure 6 shows the GLM-estimated potential elicited by the first fixation ( $\hat{f}p_{GLM}(t)$ ) for the frontal ROI in both tasks.

A clear linear trend was visible in this waveform. This trend was in line with the previous result (estimation by averaging with a common baseline), showing a quite regular increase in mean slow-wave amplitude, particularly in the RD condition. To evaluate and quantify this trend, we computed a slope. Slow-wave gradient was thus expressed in  $\mu\text{V}/\text{ms}$ , contrary to the previous analysis where the deviation was captured at each fixation rank, independently of its position in time. This computation integrated the dynamics of the fixation flow for each participant, trial and condition, taking into account the time interval between two consecutive fixations.

We performed a 2 (condition: RM, RD)  $\times$  2 (hemisphere: left, right)  $\times$  3 (ROI: frontal, central, parietal) repeated-measures ANOVA on these slope values, with  $p$  values reported after Greenhouse–Geisser correction for nonsphericity (Greenhouse and Geisser 1959) and Tukey tests for post hoc comparisons. Results revealed a Condition  $\times$  Hemisphere  $\times$  ROI interaction,  $F(2, 38) = 4.23$ ,  $p < .05$ ,  $\eta^2 = .18$ . Post hoc analysis revealed that the slopes for the frontal left ROI (Fig. 7) were significantly steeper, in terms of absolute value, in the RD ( $-.0063 \mu\text{V}/\text{ms}$ ) condition than in the RM

<sup>4</sup> For  $t > 0$  and  $< 202$  (RM) or  $169$  (RD) ms, corresponding to the mean duration of the first fixation (Table 1).





**Fig. 7** Mean slope values and standard error bars in  $\mu\text{V}/\text{ms}$  for frontal (F), central (C) and parietal (P) left and right ROI in both conditions

(.0015  $\mu\text{V}/\text{ms}$ ) condition ( $p < .05$ ). No significant differences were observed for the other ROI.

Consistent with previous analyses, results for the GLM-estimated slopes showed a larger increase in slow waves for the frontal left ROI in the RD task than in the RM one. It should be recalled that the influence of the potential evoked at text onset (Eq. 3) was eliminated from this slope computation on the estimated  $\hat{f}p_{GLM}(t)$ , as were the possible effects of oculomotor features (i.e., saccade direction and amplitude).

To go further, we took into account the fact that mean fixation duration, and consequently IFI duration, differed between tasks (significantly lower for RD than for RM), as observed in Subsection 3.2.1. Consequently, these slopes needed to be assessed not only according to task, but also as a function of IFI values, reflecting information processing

speed. Figure 8 shows the mean slope values (and mean IFI values) for each participant in each condition for the frontal left and right ROI.

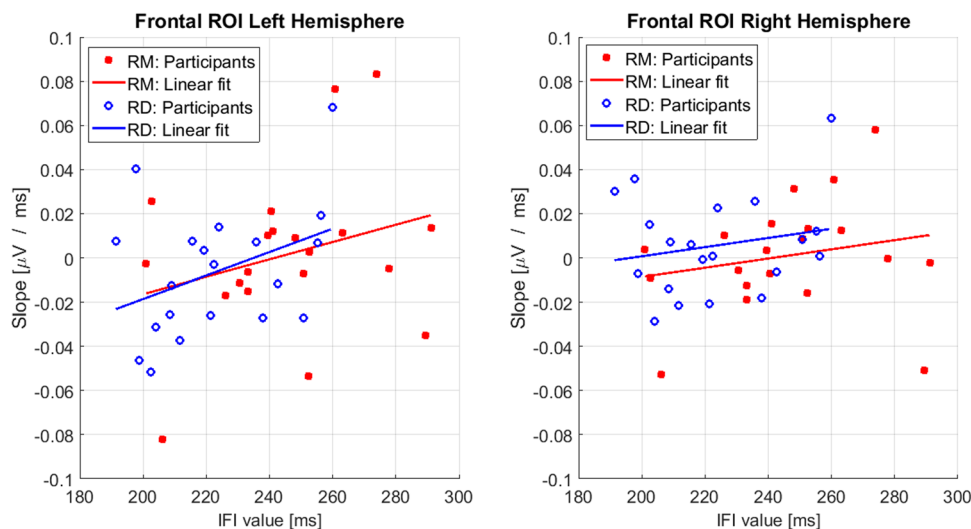
Over the left hemisphere (Fig. 8, left), linear regression analysis revealed a regression coefficient (R) of .27 for the RM task (coefficient of determination:  $R^2 = .08$ , *ns*) and .38 for the RD task ( $R^2 = .15$ ,  $p = .09$ ). Over the right hemisphere (Fig. 8, right), R was .21 for the RM task ( $R^2 = .04$ , *ns*) and .20 for the RD task ( $R^2 = .04$ , *ns*). Over the left hemisphere, R was marginally significant in the RD task, showing that the relationship between the mean slope value and IFI duration was modulated by the task.

### Discussion

The goal of the present study was to examine the effects of tasks eliciting two different ways of processing a text on EM patterns, as well as on WM involvement, investigated by EFRP. We wanted to know whether reading to memorize is more demanding for WM than reading to assess a semantic relationship with a given topic.

As well as contributing to theoretical issues (discussed below), our results shed light on major methodological issues in EFRP studies. The first methodological contribution concerns the choice of an appropriate interval for baseline correction in averaging estimation. With a *common* baseline (rest activity before text onset), our results showed that slow-wave amplitude followed a gradient between the first and fourth fixations, specifically for the frontal ROI, in line with Körner et al.'s (2014) results. Control analyses using an *individual* baseline that reset the mean value to zero at each fixation onset failed to reveal any gradient across the four fixations of interest. They confirmed that the choice of

**Fig. 8** Graphic representation of the relationship between mean IFI duration (in ms) and mean slope amplitude (in  $\mu\text{V}/\text{ms}$ ) in each task for the frontal left (left) and right (right) ROI



an appropriate baseline is very challenging in EFRP studies, and depends to a large extent on the study goals. It may generate substantially different biases in the results, as explained by Nikolaev et al. (2016).

The second contribution concerned the challenge posed by the overlap in neural responses to sequential EMs in simultaneous EEG–EM analysis. Dealing with this issue is primarily a question of deconvolving overlapping waveforms containing information that is of interest (Henderson et al. 2013). Alternative methods include the popular ADJAR algorithm (Woldorff 1993) in the context of ERP estimation, where the inter-stimulus intervals are defined in the experimental design. To our knowledge, the most appropriate method of EFRP estimation is based on GLM (for a comparison, see Kristensen et al. 2017b). The flexibility of GLM was particularly relevant in our study focusing on slow waves. Good model implementation depends mainly on two elements. The first is the definition of the potentials of interest to be investigated, which is seldom difficult. The second is the establishment of an exhaustive list of potentials that may alter the former, which is doubtless more challenging. In our study, the epoch of interest was the beginning of the task, from text onset to the end of the fourth fixation. An intuitive model for this epoch would take into account the contribution of the potential elicited at text onset, as well as those elicited at the onset of the first, second, third and fourth fixations, thus requiring five timestamps as regressors. Nevertheless, this is clearly not an appropriate model, in view of the results of averaging with an *individual* baseline, which showed that slow waves were not synchronized with particular fixations at a specific timestamp. For this reason, in our GLM configuration we used two timestamps as regressors: the timestamp at text onset to estimate the ERP elicited at the beginning of the text, and the timestamp at first fixation onset to estimate the slow waves from this event. Four supplementary timestamps at the onset of each of the first four saccades were added. As previous research had shown that saccade properties influence EFRP amplitudes and waveforms, we also included four categories of saccades in the GLM, with respect to their direction (progressive vs. regressive) and amplitude (short vs. long). This ensured that the estimated potential  $\hat{f}_{GLM}(t)$  was cleaned from the contribution of the ERP at text onset and of oculomotor parameter distortion, and provided a reliable means of observing the overall trend in neural activity on the timescale of task resolution, and more specifically in the slow waves that were generated.

Turning now to the theoretical scope of our results, behavioral and EM data clearly indicated that participants did not perform the two tasks in the same manner. Both tasks were performed correctly by participants, but the RD task seemed much easier to complete, as indicated by the success rate. Behavioral performances also showed that both the

total reading time per text and the time required to answer the question were longer in the RM condition than in the RD one. This result suggests that by the time the question was displayed, the decision-making process had already taken place in the RD task, during the reading phase.

In support of this assumption, EM results showed that the number of fixations and their duration, as well as the numbers of refixations and long progressive saccades, were higher for the RM task than for the RD task. These differences, specifically with respect to fixation duration, were both general, concerning all the fixations, and local, concerning each of the first four fixations. Such results would usually be interpreted as reflecting difficulty processing text content (Yang 2012). However, as the textual material was equivalent in the two conditions, we can assume that they reflected top-down influences on EM control. Previous results (Kaakinen and Hyönä 2010) results had indicated that task demands influence both the temporal and spatial aspects of EM, and prolonged fixations, refixation probability and saccade targeting are the result of cognitive mechanisms related to the reading intention.

In addition, reading speed was around 200 wpm for the RM task and 500 wpm for the RD task. According to Carver's (1990) classification of the different ways of reading a text, based on the reading rate (see Introduction), the RM task corresponded to *learning* (200 wpm), where readers try to acquire knowledge from the text, while the RD task corresponded to *skimming* (450 wpm in Carver's classification, 500 wpm in our study, but this difference can probably be explained by the fact that Carver's classification was based on longer texts (100–300 words) than the ones we used). *Skimming* occurs when readers need to gain a quick overview of text content. A reading mode at a slower pace is regarded as a more controlled or careful reading strategy, with more attention allocated to the task (Wotschack 2009). Overall, these results showed that readers adapt their general reading style to the task demands, and the RM task seems to need more resources than the RD task does.

As our EFRP results showed, the dependence of EM responses on task nature also applied to brain responses, as reflected by the involvement of WM, and more specifically the cognitive load induced by each task. The latter was captured through the increase in slow waves at the start of text processing, using different methodologies that all yielded similar results. In both tasks, the processing of the text elicited slow waves that gradually increased over time. Interestingly, by comparing the analyses with an *individual* vs. *common* baseline (no gradient with the *individual* baseline, as opposed to the *common* one; see above for the methodological discussion of these two choices), we were able to conclude that the slow waves were not synchronized with the local process executed at each fixation, but rather were elicited on a more global scale, depending on the nature

of the task. In other words, our result is the first, to our knowledge, to demonstrate that WM processing is continually supported by cognitive processes arising from the visual processing of information from fixations, but which are not strictly synchronized with these fixations. This result is neuropsychological evidence that WM has to fulfil two complex and alternating functions, that is, processing and storage. More specifically, WM serves to maintain reliable representations of the environment and protect them from decay and interference, at the same time turning these representations into goal-oriented behavior (Barrouillet and Camos 2015, p. 65)—though possibly not at exactly the same time, as we observed both local processes, elicited at each fixation (see averaging analyses with an *individual* baseline), and an overall process that was not time-locked to fixations.

Another primary outcome of our results is that although the behavioral and EM data suggested that the RD task was easier than the RM task, the slow waves were larger for the RD task than for the RM task, indicating that the former induced a greater memory load than the latter. These results are in line with Scharinger et al. (2015) on several points. It should be recalled that these authors observed increased load during the hyperlink-like selection processes, in terms of decreased alpha frequency band power and increased pupil dilation, compared with a baseline condition consisting of pure text reading. However, they failed to find a significant correlation between these two measures. This emphasizes the relevance of accumulating measures, as in the EFRP technique, to gain a better overview of a given cognitive process, as two measures can lead to conclusions that are not consistent, indicating that they are sensitive to different aspects of the observed process. Moreover, our results show that the RD condition may involve a concurrent task that overloads WM.

In the RM condition, participants had to retain the words in both the topic and the text, and integrate them to form the text's general meaning. This task can be regarded as a classic *storage-only* or *simple span* task, in which memory load is investigated by manipulating the number of items to be retained (e.g., Rypma et al. 1999). In the RD condition, participants had to read and constantly check whether what they had read so far was related to the topic. Processing the words and making a decision were intertwined, whereas in the RM condition, the decision was postponed until the end of reading. This assumption was confirmed by the behavioral results discussed above, as response times were shorter in the RD condition, indicating that the decision-making process had already taken place by question onset. Task switching between sentence integration and semantic comparison may require more attentional resources and therefore increase memory load (Liefvooghe et al. 2008; Vergauwe et al. 2015). The RD condition required the topic to be constantly maintained in the WM focus of attention, as it has to

be systematically compared with the words currently being processed. This is consistent with the view that the degree of control over what is in WM varies with the availability of the resources needed to exert this control (Elward et al. 2013; Elward and Wilding 2010). Recent theories and models suggest close interactions between storage and processing functions (Oberauer 2002) and between WM and attention (Awh et al. 2006), enabling readers to focus on relevant items and maintain current goals. Cowan (1995) opined that WM content can be deemed to be *activated* if representations from long-term memory are currently within the focus of attention. In this model (Cowan 1995), WM consists of the contents of short-term memory plus controlled attention. In this sense, we suggest that the RM condition mainly engaged short-term memory, whereas the RD condition was supported by WM. Along the same lines, the RD condition can be likened to the classic *complex span* task, which can be described as a dual task requiring simultaneous short-term retention of some information and processing of other information. Complex span tasks require items to be maintained even though other information needs to be processed, and may reflect the ability to resist interference. This view can be summarized by the equation  $\text{complex span} = \text{simple span} + \text{controlled attention}$  (Oberauer et al. 2004).

Another possible explanation for the EEG differences between the two tasks, as well as the lack of consistency between EM and EEG results, is the difference in time pressure. We wanted to construct tasks that were as ecological as possible, and the time pressure in the RD task was probably typical of online web reading, where a huge number of web pages addressing a given topic are available and have to be skimmed for relevance (Scharinger et al. 2015). The first indicator of the impact of time pressure was the latency of the first fixation after text onset, which was shorter for the RD task than for the RM task. This showed that participants engaged in different mental preparation, depending on the task. The skimming observed in the RD task is a reading strategy that is classically observed (see literature on web browsing) when readers have to cope with too much text to read in the time available to them (Duggan and Payne 2009). Finally, WM is generally considered to be correlated with task demands as well as with time pressure (Paas et al. 1994). Studies have shown that a higher level of arousal and/or mental workload under time pressure modulates EEG within the alpha and/or gamma frequency bands over frontocentral areas (Cheng 2018), which could also explain the greater slow wave increase for our RD task.

This second interpretation does not exclude the first one, as the WM time-based resource sharing (TBRS) model (Barrouillet et al. 2004, 2007) assumes that cognitive load can be estimated as the proportion of time devoted to extra tasks that impede the processes currently taking place, and that items need to be refreshed to avoid their decay over

time. This model assumes that both processing and storage require attention, viewed as a pool of limited domain-general resources that have to be shared (Vergauwe et al. 2015; see also Kiyonaga and Egner 2014a, b, for a similar assumption). Our results are in line with this theoretical point of view, albeit using a completely different and more ecological measure of memory load. In the RD condition, the topic with which the texts had to be compared, as well as the semantic comparisons that recurred throughout a given trial, meant extra information to manage that increased the proportion of time devoted to what could be viewed as distracting activities. Accordingly, the time taken to perform these extra activities in the RD condition may have generated a need for more attentional resources to compensate for the decay of information that had to be maintained throughout the trial.

Furthermore, our results showed that the increase in slow waves was mainly left lateralized. Even if EEG is known as a brain imaging technique with particularly poor spatial resolution (Burle et al. 2015), this result is consistent with research showing left hemispheric lateralization for verbal WM, most notably in the frontal and parietal lobes, in contrast to more right hemispheric lateralization for spatial WM (Nagel et al. 2013). Our results are in line with studies reporting left-lateralized alpha frequency band power effects for linguistic task material (Spironelli and Angrilli 2010; Strauß et al. 2014). The literature on this topic is still inconclusive, but our data point to modality-specific WM (i.e., verbal and visuospatial WM) with different lateralization patterns according to the material to be maintained (Golby et al. 2001).

Finally, both behavioral and EM results showed that the visual acquisition rate, that is, the speed at which visual acquisition occurs, was faster in the RD condition. More specifically, our results showed that in the RD task, the faster the fixation rate (indexed by IFI duration), the steeper the slow wave slope. This was in line with experiments showing that the mental load generated by a task depends on the pace at which distractors are processed, and is not necessarily determined by their number (Barrouillet et al. 2011, 2013). In our experiment, the reading process could be regarded as the *main task*, and the decision-making processes as *distractors*. Temporal factors also need to be taken into account to understand better how WM operates.

One limitation of the present research that has to be addressed is that we decided to administer the two tasks in the same order throughout (RM then RD), in order to avoid the influence of the RD task on the RM one. We wanted the two tasks to be as similar as possible, except that in one task, participants had to read and memorize, whereas in the other, they had to read and decide. We were concerned that if the RD task was presented before the RM, participants would be inclined to engage in a decision-making process in the RM task too. However,

we do not think that this choice generated confound effects (specifically fatigue) for several reasons. First, when we performed pilot tests, none of the participants reported any accumulated fatigue. Second, as reported above, behavioral results for the RD task were almost perfect (97.8% correct answers), even though mental fatigue is traditionally associated with a decrease in task performance (Xie et al. 2016). More specifically, as regards the literature on the speed/accuracy trade-off, fatigue can lead to the deterioration of one or the other (i.e., impaired performance speed, but intact accuracy, or vice versa), or else have a detrimental effect on both speed and accuracy (Alhola and Polo-Kantola 2007). Again, in our case, results did not show any impairment of response speed in the RD task (< 1 s). Finally, the many studies that have reported fatigue-related deficits in attentional processing and/or memory functions tested cognitive performance in tiring conditions that were much more demanding than ours, specifically in sleep deprivation conditions (Lowe et al. 2017).

Notwithstanding this limitation, we can conclude that coregistering EEG and EMs is a more powerful method than carrying out separate registration to understand the processes involved in an unconstrained visual paradigm. This method had limitations for which our study proposed solutions and precautions. We hope that the latter will make EFRP simpler and encourage its more frequent use.

Although the two tasks were part of a controlled laboratory experiment, they are typical of activities we all perform on a daily basis. Investigating the reading process is crucial to theories of reading, but our study also improves understanding of how different environments affect and constrain this process (Simola 2011). In particular, owing to the huge amount of data that are now available, searching for information on the Internet has become a significant activity in people's daily lives. Our study enhances the understanding of this new form of reading, as a result of the Internet's increasing dominance as a major source of information. This activity differs from classic reading because the extra information being searched for has to remain active in WM. Our study shows that, in contrast to intuition or some forms of work (Wotschack 2009), this has a higher cognitive cost that is visible in EEG signals when they are jointly recorded with EM.

**Acknowledgements** The authors thank the “Délégation à la Recherche Clinique et à l’Innovation” of Grenoble “Centre Hospitalier Universitaire” (CHU) for its role in the ethics committee, particularly Beatrice Portal and Dominique Garin. EEG/eye-tracker co-registration was performed at the IRMaGe Neurophysiology facility in Grenoble (France), which was partly funded by the French program “Investissement d’Avenir” run by the “Agence Nationale pour la Recherche” (grant ‘Infrastructure d’Avenir en Biologie Santé’ - ANR-11-INBS-0006). The present study was part-funded by grants from the “Pôle Grenoble Cognition” (PGC\_AAP2014), and by a grant from the LabEx PERSYVAL-Lab (ANR-11-LABX-0025-01).

## Appendix 1: EFRP by Averaging with an *Individual* Baseline

Results from the ANOVA revealed the following significant effects:

Condition:  $F(1, 19) = 23.16, p < .001, \eta^2 = .55$

Condition  $\times$  Hemisphere interaction:  $F(1, 19) = 4.74, p < .05, \eta^2 = .20$

Condition  $\times$  Hemisphere  $\times$  ROI interaction:  $F(2, 38) = 7.62, p = .01, \eta^2 = .29$

Condition  $\times$  Hemisphere  $\times$  Rank interaction:  $F(3, 57) = 3.20, p < .05, \eta^2 = .14$

Condition  $\times$  ROI  $\times$  Rank interaction:  $F(6, 114) = 3.56, p < .05, \eta^2 = .16$

Condition  $\times$  Hemisphere  $\times$  ROI  $\times$  Rank interaction:  $F(6, 114) = 2.98, p < .05, \eta^2 = .14$

Table 3 sets out the results of the post hoc comparison for the 4-factor interaction (i.e., Condition  $\times$  Hemisphere  $\times$  ROI  $\times$  Rank), for frontal, central and parietal ROI.

Figure 9 illustrates the EFRP results with an individual baseline for the frontal ROI (i.e., F4, F6, F8 and FC6 for the right hemisphere and F3, F5, F7 and FC5 for the left one).

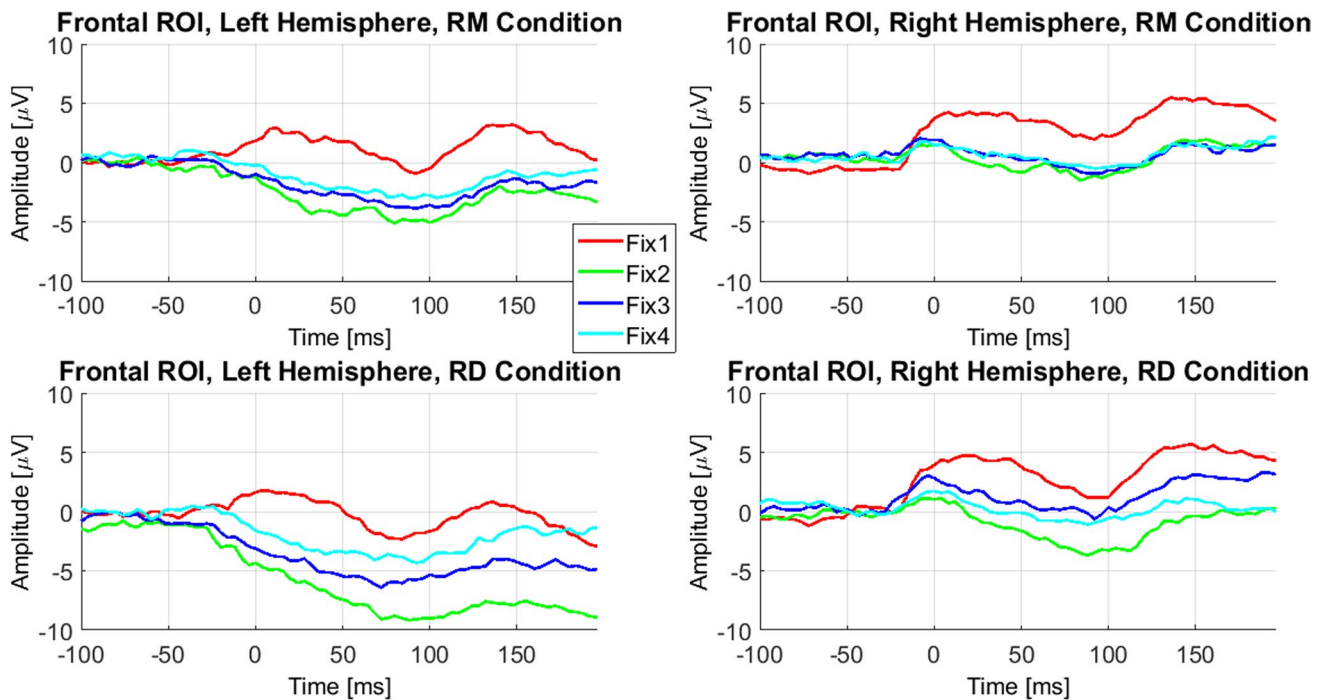
As emphasized in Subsection 3.3.2 (“Estimation by averaging”), an *individual* baseline resets the activity to zero (Körner et al. 2014) at each fixation onset. For example, the mean activity of the EFRP baseline at fixation  $n$  is roughly equivalent to the mean activity elicited at fixation  $n - 1$ . Our results showed no major increment or decrement in synchronized activity elicited at one fixation compared with the following fixation, showing that the level of synchronized activity at fixation rank  $n$  was roughly the same as that at fixation rank  $n + 1$ . As expected, slow-wave activity was not synchronized with each fixation. There were several significant differences between the two tasks, with larger mean amplitudes for the RD task than for the RM one, showing that the contribution of a specific fixation to slow-wave amplitude was greater in the RD condition than in the RM one.

**Table 3** Mean (standard error) amplitude values, in  $\mu\text{V}$ , in the latency ranges of interest for frontal, central and parietal ROI in both conditions with an individual baseline

		Rank 1	Rank 2	Rank 3	Rank 4	Significant differences
Frontal ROI amplitude [ $\mu\text{V}$ ]						
Left Hem.	RM	1.52 (.73)	-3.69 (.69)	-2.69 (.67)	-1.98 (.48)	1 > (4=3) > 2***
Left Hem.	RD	-.10 (.74)	-7.50 (.95)	-4.91 (.72)	-3.07 (.72)	1 > 4 > 3 > 2***
Significant differences		RM > RD ***	RM > RD ***	RM > RD ***	RM > RD ***	
Right Hem.	RM	3.70 (.83)	-.03 (.52)	.37 (.52)	.49 (.56)	1 > 2 = 3 = 4***
Right Hem.	RD	3.60 (.89)	-1.63 (.64)	1.18 (.51)	.06 (.87)	1 > 3 > 4 > 2***
Significant differences		ns	RM > RD ***	RM > RD *	ns	
Central ROI amplitude [ $\mu\text{V}$ ]						
Left Hem.	RM	3.15 (.76)	-2.45 (.60)	-1.71 (.56)	-1.19 (.43)	1 > (4=3=2) & 4 > 2***
Left Hem.	RD	2.18 (.68)	-5.17 (.68)	-3.14 (.48)	-2.31 (.63)	1 > 4 > 3 > 2***
Significant differences		RM > RD ***	RM > RD ***	RM > RD ***	RM > RD ***	
Right Hem.	RM	4.05 (.82)	-.98 (.53)	-.68 (.50)	-.10 (.50)	1 > (2=3=4) & 4 > 2**
Right Hem.	RD	3.13 (.90)	-2.69 (.58)	-.78 (.39)	-1.07 (.72)	1 > 2 > 3 = 4***
Significant differences		RM > RD **	RM > RD ***	ns	RM > RD ***	
Parietal ROI amplitude [ $\mu\text{V}$ ]						
Left Hem.	RM	4.64 (.80)	-2.49 (.62)	-1.73 (.54)	-.64 (.44)	1 > (2=3) > 4***
Left Hem.	RD	4.16 (.85)	-4.28 (.57)	-2.89 (.53)	-2.14 (.65)	1 > (3=4) > 2***
Significant differences		ns	RM > RD ***	RM > RD ***	RM > RD ***	
Right Hem.	RM	5.00 (.88)	-2.09 (.62)	-1.56 (.54)	-.43 (.49)	1 > 4 > (2=3)***
Right Hem.	RD	4.25 (1.02)	-3.71 (.58)	-2.23 (.49)	-1.69 (.67)	1 > (3=4) > 2***
Significant differences		ns	RM > RD ***	ns	RM > RD ***	

Significant differences in rows correspond to post hoc analyses between the four fixation ranks, and those in columns correspond to post hoc analyses between the two tasks. \* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$





**Fig. 9** Grand-average EFRP, estimated by averaging with an individual baseline, for the RM task (top) and RD task (bottom), for the first four fixations on the text, for the frontal right (right) and left (left) ROI

## Appendix 2: Implementation and Configuration of the GLM

In the present study, GLM were used to estimate the neural activity time-locked to the period between first fixation onset and fourth fixation offset. This activity can be written as  $fp(t)$ . The latency of the first fixation onset was around 200 ms (see Table 4), showing that, as previously emphasized, this activity overlapped with the potential elicited at text onset  $s(t)$ .

These two potentials were included in the model as described by Eq. 3 (see Subsection 3.4.1), with the potential evoked by the first fixation plus the potentials elicited by the successive fixations in the selected time window. The saccadic potentials elicited at the four first saccade onsets were also considered in the model, as the distribution of the incoming saccades, in terms of direction and amplitude, differed across conditions (see Subsection 3.2.2). Let us recall here the equation  $x_i(t) = s(t) + fp(t - \tau_i^{(1)}) + \sum_{c=1}^4 sp_c(t - \tau_{c,i}^{(l)}) + n_i(t)$ , where  $x_i(t)$  is the signal time-locked to text onset during the  $i$ th epoch,  $s(t)$  the potential elicited at text onset,  $fp(t)$  the potential evoked at first fixation onset plus the potentials elicited at subsequent fixation onsets up to the end of the epoch,  $sp_c(t)$  the saccadic potential elicited at saccade onset for a given category ( $c$ ),  $\tau_i^{(1)}$  the

timestamp of the first fixation onset,  $\tau_{c,i}^{(l)}$  the timestamp of the  $l$ th saccade onset for the category  $c$ , and  $n_i(t)$  the noise of the ongoing activity. We considered four categories of saccades: progressive vs. regressive, and short vs. long.

This appendix shows the mathematical implementation leading to the final estimation. To estimate  $s(t)$  and  $fp(t)$  by ordinary least square regression, Eq. 3 can be rewritten in matrix form:

$$\forall i \in \{1, \dots, E\}, \quad \mathbf{x}_i = \mathbf{D}_s \cdot \mathbf{s} + \mathbf{D}_{fp,i} \cdot \mathbf{fp} + \sum_{c=1}^4 \mathbf{D}_{sp,c,i} \cdot \mathbf{sp}_c + \mathbf{n}_i \quad (4)$$

$\mathbf{x}_i$  is the vector ( $\mathbf{x}_i = [x_i(1), \dots, x_i(N_e)]^\dagger$ ;  $x_i \in \mathbb{R}^{N_e}$ ) of the observed EEG samples time-locked to text onset, for the  $i$ th epoch, with  $[\cdot]^\dagger$  the transpose operator.  $N_e$  is the number of samples, that is, the length of the observed signal  $x_i(t)$ .  $\mathbf{n}_i$  is the noise vector ( $\mathbf{n}_i = [n_i(1), \dots, n_i(N_e)]^\dagger$ ;  $\mathbf{n}_i \in \mathbb{R}^{N_e}$ ).

**Table 4** Statistical summary of the latency (in ms) of the first and fourth fixation onsets based on individual means

Condition	Mean first fixation latency (SE)	
	RM	RD
Fix. 1	206.82 (4.78)	195.77 (4.10)
Fix. 4	975.18 (20.73)	886.31 (18.40)

with the same number of samples.  $\mathbf{s} \in \mathbb{R}^{N_s}$  is the vector of the response time-locked to text onset, and  $N_s$  is the length of the response  $s(t)$ .  $\mathbf{fp} \in \mathbb{R}^{N_{fp}}$  is the vector of the response time-locked to first fixation onset and  $N_{fp}$  is the length of the response  $fp(t)$ . For a given saccade category  $c$ ,  $\mathbf{sp}_c \in \mathbb{R}^{N_{sp}}$  is the vector of the saccadic response time-locked to saccade onset and  $N_{sp}$  is the length of the response  $sp_c(t)$ .  $\mathbf{D}_s \in \mathbb{R}^{N_e \times N_s}$  is the Toeplitz matrix<sup>5</sup> with  $N_e$  rows and  $N_s$  columns, for text onset.  $\mathbf{D}_s$  is defined by its first column, with entries that are all equal to zero except for one at the row subscript corresponding to the 0 ms position in the epoch.  $\mathbf{D}_{fp,i} \in \mathbb{R}^{N_e \times N_{fp}}$  is the Toeplitz matrix with  $N_e$  rows and  $N_{fp}$  columns, for coding the first fixation onset during the  $i$ th epoch.  $\mathbf{D}_{fp,i}$  is defined by its first column, with entries that are all equal to zero except one, at the row subscript corresponding to the  $\tau_i^{(1)}$  ms position in the  $i$ th epoch. The Toeplitz matrices  $\mathbf{D}_{sp,c,i}$  ( $c = 1 \dots 4$ ),  $N_e$  rows and  $N_{sp}$  columns, were set up in the same way, but using the timestamps  $\tau_{c,i}^{(l)}$  for the row subscripts. Unlike the matrices  $\mathbf{D}_{fp,i}$  and  $\mathbf{D}_{sp,c,i}$ , the  $\mathbf{D}_s$  matrix does not depend on the epoch number ( $i$ ), as the timestamps of text onset are always equal to zero whatever the epoch.  $\mathbf{D}_s$ ,  $\mathbf{D}_{fp,i}$  and  $\mathbf{D}_{sp,c,i}$  ( $c = 1..4$ ) are sparse matrices. Matrices  $\mathbf{D}_s$  and  $\mathbf{D}_{fp,i}$  are composed of only one diagonal equal to one, all other values being equal to zero. For the matrices  $\mathbf{D}_{sp,c,i}$  ( $c = 1..4$ ), zero, one, two, three or four diagonals are equal to one, depending on the number of saccades belonging to a given category ( $c$ ), and all other values are equal to zero. Considering all epochs ( $E$ ), the observations are concatenated such that:

$$\mathbf{x} = \mathbf{D}_s \mathbf{s} + \mathbf{D}_{fp} \mathbf{fp} + \sum_{c=1}^4 \mathbf{D}_{sp,c} \mathbf{sp}_c + \mathbf{n} \quad (5)$$

with  $\mathbf{x} = [\mathbf{x}_1^\dagger, \dots, \mathbf{x}_E^\dagger]^\dagger \in \mathbb{R}^N$ ,  $\mathbf{D}_s = [\mathbf{D}_s^\dagger, \dots, \mathbf{D}_s^\dagger]^\dagger \in \mathbb{R}^{N \times N_s}$ ,  $\mathbf{D}_{fp} = [\mathbf{D}_{fp,1}^\dagger, \dots, \mathbf{D}_{fp,E}^\dagger]^\dagger \in \mathbb{R}^{N \times N_{fp}}$  and  $\mathbf{D}_{sp,c} = [\mathbf{D}_{sp,c,1}^\dagger, \dots, \mathbf{D}_{sp,c,E}^\dagger]^\dagger \in \mathbb{R}^{N \times N_{sp}}$  ( $c = 1..4$ ), where  $N = N_e \times E$ ,  $N$  is the total number of samples. After concatenation of the Toeplitz matrices and evoked potentials, Eq. 6 becomes  $\mathbf{x} = \mathbf{D} \mathbf{p} + \mathbf{n}$  with  $\mathbf{D}$  equal to  $[\mathbf{D}_s, \mathbf{D}_{fp}, \mathbf{D}_{sp,1}, \dots, \mathbf{D}_{sp,4}]$  and  $\mathbf{p}$  is the concatenation of the six evoked potentials such that  $\mathbf{p} = [\mathbf{s}^\dagger, \mathbf{fp}^\dagger, \mathbf{sp}_1^\dagger, \dots, \mathbf{sp}_4^\dagger]^\dagger$ . The solution given by the least square minimization is:

$$\hat{\mathbf{p}}_{GLM} = (\mathbf{D}^\dagger \mathbf{D})^{-1} \mathbf{D}^\dagger \mathbf{x} \quad (6)$$

<sup>5</sup> By definition, a Toeplitz matrix is a descending diagonal-constant matrix.

where  $\hat{\mathbf{p}}_{GLM}$  is the concatenation of both estimates, such that  $\hat{\mathbf{p}}_{GLM} = [\hat{\mathbf{s}}_{GLM}^\dagger, \hat{\mathbf{fp}}_{GLM}^\dagger, \hat{\mathbf{sp}}_{1GLM}^\dagger, \dots, \hat{\mathbf{sp}}_{4GLM}^\dagger]^\dagger$ .

As for the estimation by averaging, the model was applied separately for each participant. The grand average was then obtained by averaging all estimates for all participants.

Here, the main configuration parameters for the GLM are (1) the time intervals for estimated signals and (2) the time interval of the epoch for the observed signal  $x(t)$ , providing the number of samples<sup>6</sup> (i.e.  $N_s$ ,  $N_{fp}$  and  $N_e$ ) for the evoked potential  $s(t)$ ,  $fp(t)$  and observed signal  $x(t)$ , respectively. The estimation window of the potential  $s(t)$  extended from 100 ms (baseline computation between  $-100$  and  $0$  ms) before text onset to 700 ms<sup>7</sup> after. Thus, the total duration for the potential  $s(t)$  was 800 ms, thereby defining the number  $N_s$  of samples. The estimation window of the potential  $fp(t)$  extended from 200 ms (baseline computation between  $-200$  and  $-100$  ms) before first fixation onset to 920 ms after, for the RM condition (840 ms after for the RD condition). This time value was set to capture the first four fixation onsets, that is, the mean IFI between the first and fourth fixations (see Table 4) was 768 ms, rounded up to  $770 + 150$  ms ( $690 + 150$  ms for the RD condition). The total duration for the potential  $fp(t)$  was 1120 ms for the RM condition (1050 ms for the RD condition), thereby defining the number  $N_{fp}$  of samples. The estimation window for the saccadic potential  $sp_c(t)$  ( $c = 1..4$ ) was configured from 50 ms (baseline computation between  $-50$  and  $-10$  ms) before saccade onset to 200 ms after. The epoch duration for the observed signal  $x(t)$  was set from 100 ms before text onset to 1150 ms after for the RM condition (1050 ms for the RD condition). These values were chosen as a function of the latency of the fourth fixation ( $+150$  ms). The number  $N_e$  of samples of the observed signals was then defined.

## References

- Alhola P, Polo-Kantola P (2007) Sleep deprivation: impact on cognitive performance. *Neuropsychiatric Dis Treat* 3(5):553–567
- Awh E, Vogel EK, Oh S-H (2006) Interactions between attention and working memory. *Neurosciences* 139:201–208
- Baggio G, Hagoort P (2011) The balance between memory and unification in semantics: towards a dynamic account of the N400. *Lang Cogn Process* 26:1338–1367
- Bardy F, Dillon H, Dun BV (2014) Least-squares deconvolution of evoked potentials and sequence optimization for multiple stimuli under low-jitter conditions. *Clin Neurophysiol* 125:727–737
- Barrouillet P, Camos V (2015) Working memory: loss and reconstruction. Psychology Press, Hove

<sup>6</sup> The number of samples is the duration of the time interval multiplied by the sampling frequency.

<sup>7</sup> 700 ms is the usual duration for this transient response.

- Barrouillet P, Bernardin S, Camos V (2004) Time constraints and resource-sharing in adults' working memory spans. *J Exp Psychol Gen* 133:83–100
- Barrouillet P, Bernardin S, Portrat S, Vergauwe E, Camos V (2007) Time and cognitive load in working memory. *J Exp Psychol* 33(3):570–585
- Barrouillet P, Portrat S, Camos V (2011) On the law relating processing to storage in working memory. *Psychol Rev* 118:175–192
- Barrouillet P, Plancher G, Guida A, Camos V (2013) Forgetting at short term: when do event-based interference and temporal factors have an effect? *Acta Physiol (Oxford)* 142:155–167
- Bastiaansen MCM, Hagoort P (2015) Frequency-based segregation of syntactic and semantic unification during online sentence level language comprehension. *J Cogn Neurosci* 27(11):2095–2107
- Burle B, Spieser L, Roger C, Casini L, Hasbroucq T, Vidal F (2015) Spatial and temporal resolutions of EEG: is it really black and white? a scalp current density view. *Int J Psychophysiol* 97(3):210–220
- Burns MD, Bigdely-Shamlo N, Smith NJ, Kreutz-Delgado K, Makeig S (2013) Comparison of averaging and regression techniques for estimating event related potentials. In *IEEE engineering in biology and medicine conference, Osaka*, pp 1680–1683
- Carver R (1990) *Reading rate: a review of research and theory*. Academic Press, New York
- Cheng SY (2018) Evaluation of effect on cognition response to time pressure by using EEG. In: Duffy V, Lightner N (eds) *Advances in human factors and ergonomics in healthcare and medical devices. AHFE 2017. Advances in intelligent systems and computing*, vol 590. Springer, Cham
- Congedo M, Korczowski L, Delorme A, da Silva FL (2016) Spatio-temporal common pattern: a companion method for ERP analysis in the time domain. *J Neurosci Methods* 267:74–88
- Cowan N (1995) *Attention and memory: an integrated framework*. Oxford Psychology Series, No. 26. Oxford University Press, New York
- Dandekar S, Privitera C, Carney T, Klein SA (2012) Neural saccadic response estimation during natural viewing. *J Neurophysiol* 107:1776–1790
- Dimigen O, Sommer W, Hohlfield A, Jacobs AM, Kliegl R (2011) Coregistration of eye movements and EEG in natural reading: analysis and review. *J Exp Psychol Gen* 140:552–572
- Ditman T, Holcomb PJ, Kuperberg GR (2007) An investigation of concurrent ERP and self-paced reading methodologies. *Psychophysiology* 44(6):927–935
- Duggan GB, Payne SJ (2009) Text skimming: the process and effectiveness of foraging through text under time pressure. *J Exp Psychol Appl* 15(3):228–242
- Elward RL, Wilding EL (2010) Working memory capacity is related to variations in the magnitude of an electrophysiological marker of recollection. *Brain Res* 1342:55–62
- Elward RL, Evans LH, Wilding EL (2013) The role of working memory capacity in the control of recollection. *Cortex* 49:1452–1462
- Frey A, Ionescu G, Lemaire B, Lopez-Orozco F, Baccino T, Guérin-Dugué A (2013) Decision-making in information seeking on texts: an eye-fixation-related potentials investigation. *Front Neurosci (Syst Neurosci)* 7:39
- Golby AJ, Poldrack RA, Brewer JB, Spencer D, Desmond JE, Aron AP, Gabrieli JD (2001) Material-specific lateralization in the medial temporal lobe and prefrontal cortex during memory encoding. *Brain* 124:1841–1854
- Gordon M, Pathak P (1999) Finding information on the world wide web: the retrieval effectiveness of search engines. *Inf Process Manag* 35(2):141–180
- Gordon PC, Hendrick R, Levine WH (2002) Memory-load interference in syntactic processing. *Psychol Sci* 13:425–430
- Greenhouse S, Geisser S (1959) On methods in the analysis of profile data. *Psychometrika* 24:95–112
- Hagoort P (2013) MUC (memory, unification, control) and beyond. *Front Psychol* 4:416
- Hagoort P (2016) MUC (memory, unification, control): a model on the neurobiology of language beyond single word processing. In: Hickok G, Small S (eds) *Neurobiology of language*, vol 28. Elsevier, Amsterdam, pp 339–347
- Henderson JM, Luke SG, Schmidt J, Richards JE (2013) Co-registration of eye movements and event-related potentials in connected-text paragraph reading. *Front Syst Neurosci* 7:28
- Hutzel F, Braun M, Vö M-L, Engl V, Hofmann M, Dambacher M, Leder H, Jacobs AM (2007) Welcome to the real world: validating fixation-related brain potentials for ecologically valid settings. *Brain Res* 1172:124–129
- Ionescu G, Guyader N, Guérin-Dugué A (2009) SoftEye software. Registration number: IDDNFR001.200017.000.S.P.2010.003.31235
- Jasper H (1958) The ten-twenty electrode system of the International Federation. *Electroencephalogr Clin Neurophysiol* 10:371–375
- Jung TP, Makeig S, Humphries C, Lee TW, McKeown MJ, Iragui V, Sejnowski TJ (2000) Removing electroencephalographic artifacts by blind source separation. *Psychophysiology* 37(2):163–178
- Kaakinen JK, Hyönä J (2010) Task effects on eye movements during reading. *J Exp Psychol* 36:1561–1566
- Khader P, Heil M, Rösler F (2005) Material-specific long-term memory representations of faces and spatial positions: evidence from slow event-related brain potentials. *Neuropsychologia* 43:2109–2124
- Kintsch W (1988) The use of knowledge in discourse processing: a construction-integration model. *Psychol Rev* 95:163–182
- Kintsch W (1998) *Comprehension: a paradigm for cognition*. Cambridge University, New York
- Kiyonaga A, Egner T (2014a) The working memory stroop effect: when internal representations clash with external stimuli. *Psychol Sci* 25(8):1619–1629
- Kiyonaga A, Egner T (2014b) Resource-sharing between internal maintenance and external selection modulates attentional capture by working memory content. *Front Hum Neurosci* 8:670
- Kliegl R, Dambacher M, Dimigen O, Jacobs AM, Sommer W (2012) Eye movements and brain electric potentials during reading. *Psychol Res* 76(2):145–158
- Körner C, Braunstein V, Stangl M, Schlögl A, Neuper C, Ischebeck A (2014) Sequential effects in continued visual search: using fixation related potentials to compare distractor processing before and after target detection. *Psychophysiology* 51:385–395
- Kristensen E, Guérin-Dugué A, Rivet B (2017a) Regularization and a general linear model for event-related potential estimation. *Behav Res Methods* 8:1–20
- Kristensen E, Rivet B, Guérin-Dugué A (2017b) Estimation of overlapped eye fixation related potentials: the General Linear Model, a more flexible framework than the ADJAR algorithm. *J Eye Mov Res* 10(1):1–27
- Landauer TK, McNamara DS, Dennis S, Kintsch W (eds) (2007) *Handbook of latent semantic analysis*. Erlbaum, Mahwah
- Lemaire B, Denhière G, Bellissens C, Jhean-Larose S (2006) A computational model for simulating text comprehension. *Behav Res Methods* 38(4):628–637
- Leu DL, Forzani E, Rhoads C, Maykel C, Kennedy C, Timbrell N (2015) The new literacies of online research and comprehension: rethinking the reading achievement gap. *Read Res Q* 50(1):37–59
- Liefoghe B, Barrouillet P, Vandierendonck A, Camos V (2008) Working memory costs of task switching. *J Exp Psychol* 34(3):478–494
- Lowe CJ, Safati A, Hall PA (2017) The neurocognitive consequences of sleep restriction: a meta-analytic review. *Neurosci Biobehav Rev* 80:586–604
- Luck S (2005) *An introduction to the event-related potential technique*. MIT Press, Cambridge

- Makeig S, Bell AJ, Jung T-P, Ghahremani D, Sejnowski TJ (1997) Blind separation of auditory event-related brain responses into independent components. *Proc Natl Acad Sci USA* 94:10979–10984
- Missonnier P, Leonards U, Gold G, Palix J, Ibanez V, Giannakopoulos P (2003) A new electrophysiological index for working memory load in humans. *Neuroreport* 14:1451–1455
- Montfort V, Pouthas V (2003) Effects of working memory demands on frontal slow waves in time-interval reproduction tasks in humans. *Neurosci Lett* 343:195–199
- Nagel BJ, Herting MM, Maxwell EC, Bruno R, Fair D (2013) Hemispheric lateralization of verbal and spatial working memory during adolescence. *Brain cognition* 82(1):58–68
- Nikolaev AR, Pannasch S, Ito J, Belopolsky AV (2014) Eye movement-related brain activity during perceptual and cognitive processing. *Front Syst Neurosci* 8:62
- Nikolaev AR, Meghanathan RN, van Leeuwen C (2016) Combining EEG and eye movement recording in free viewing: pitfalls and possibilities. *Brain Cogn* 107:55–83
- Oberauer K (2002) Access to information in working memory: exploring the focus of attention. *J Exp Psychol* 28(3):411–421
- Oberauer K, Lange E, Engle RW (2004) Working memory capacity and resistance to interference. *J Mem Lang* 51:80–96
- Ossandón JP, Helo AV, Montefusco-Siegmund R, Maldonado PE (2010) Superposition model predicts EEG occipital activity during free viewing of natural scenes. *J Neurosci* 30(13):4787–4795
- Paas FG, Van Merriënboer JJ, Adam JJ (1994) Measurement of cognitive load in instructional research. *Percept Mot Skills* 79:419–430
- Pan B, Hembrooke H, Joachims T, Lorigo L, Gay G, Granka L (2007) In google we trust: users' decisions on rank, position, and relevance. *J Comput-Mediat Commun* 12(3):801–823
- Radach R, Huestegge L, Reilly R (2008) The role of top down factors in local eye movement control during reading. *Psychol Res* 72:675–688
- Reichle ED, Reingold EM (2013) Neurophysiological constraints on the eye-mind link. *Front Hum Neurosci* 7:361
- Rösler F (1993) Beyond reaction time and error rate: monitoring mental processes by means of slow event-related brain potentials. In: McCallum WC, Curry SH (eds) *Slow potential changes in the human brain*. Plenum Press, New York, pp 105–119
- Ruchkin DS (1965) An analysis of average response computations based upon aperiodic stimuli. *IEEE Trans Biomed Eng* 12(2):87–94
- Rypma B, Prabhakaran V, Desmond JE, Glover GH, Gabrieli JD (1999) Load-dependent roles of frontal brain regions in the maintenance of working memory. *Neuroimage* 9:216–226
- Scharinger C, Kammerer Y, Gerjets P (2015) Pupil dilatation and EEG alpha frequency band power reveal load on executive functions for link-selection processes during text reading. *PLoS ONE* 10(6):1–24
- Simola J (2011) Investigating online reading with eye tracking and EEG: the influence of text format, reading task and parafoveal stimuli on reading processes. Academic dissertation, University of Helsinki, Institute of Behavioural Sciences, Studies in Psychology, 81:2011
- Simola J, Salojärvi J, Kojo I (2008) Using hidden Markov model to uncover processing states from eye movements in information search tasks. *Cogn Syst Res* 9(4):237–251
- Simola J, Holmqvist K, Lindgren M (2009) Right visual field advantage in parafoveal processing: evidence from eye-fixation-related potentials. *Brain Lang* 111(2):101–113
- Simola J, Torniaainen J, Moisala M, Kivikangas M, Krause CM (2013) Eye movement related brain responses to emotional scenes during free viewing. *Front Syst Neurosci* 7(41):1–16
- Simola J, LeFevre K, Torniaainen J, Baccino T (2015) Affective processing in natural scene viewing: valence and arousal interactions in eye-fixation-related potentials. *Neuroimage* 106:21–33
- Smith NJ, Kutas M (2015a) Regression-based estimation of ERP waveforms: I. The rERP framework. *Psychophysiology* 52(2):157–168
- Smith NJ, Kutas M (2015b) Regression-based estimation of ERP waveforms: II. Non-linear effects, overlap correction, and practical considerations. *Psychophysiology* 52(2):169–181
- Spironelli C, Angrilli A (2010) Developmental aspects of language lateralization in delta, theta, alpha and beta EEG bands. *Biol Psychol* 85:258–267
- Strauß A, Kotz S, Scharinger M, Obleser J (2014) Alpha and theta brain oscillations index dissociable processes in spoken word recognition. *Neuroimage* 97:387–395
- Underwood G, Radach R (1998) Eye guidance and visual information processing: reading, visual search, picture perception and driving. In: Underwood G (ed) *Eye guidance in reading and scene perception*. Elsevier, Oxford, pp 1–27
- van Vugt MK, Simen P, Nystrom L, Holmes P, Cohen JD (2014) Lateralized readiness potentials reveal properties of a neural mechanism for implementing a decision threshold. *PLoS ONE*, 9(3):e90943
- Vergauwe E, Hartstra E, Barrouillet P, Brass M (2015) Domain-general involvement of the posterior frontolateral cortex in time-based resource-sharing in working memory: an fMRI study. *NeuroImage* 115:104–116
- Weaver CA III, Mannes S, Fletcher CR (2012) *Discourse comprehension: essays in Honor of Walter Kintsch*. Lawrence Erlbaum Associates, Hillsdale
- Weger UW, Inhoff AW (2007) Long-range regressions to previously read words are guided by spatial and verbal memory. *Mem Cogn* 35(6):1293–1306
- Woldorff MG (1993) Distortion of ERP averages due to overlap from temporally adjacent ERPs: analysis and correction. *Psychophysiology* 30:98–119
- Woodman GF (2010) A brief introduction to the use of event-related potentials (ERPs) in studies of perception and attention. *Atten Percept Psychophys* 72(8):2031–2046
- Wotschack C (2009) Eye movements in reading strategies: how reading strategies modulate effects of distributed processing and oculomotor control. Doctoral dissertation, Universität Potsdam, Potsdam, Germany
- Xie J, Xu G, Wang J, Li M, Han C, Jia Y (2016) Effects of mental load and fatigue on steady-state evoked potential based brain computer interface tasks: a comparison of periodic flickering and motion-reversal based visual attention. *PLoS ONE* 11(9):e0163426
- Yang S (2012) Effects of processing difficulty on eye movements in reading: a review of behavioral and neural observations. *J Eye Mov Res* 5(4):1, 1–16
- Yarbus AL (1967) *Eye movements and vision*. Plenum, New York